RETURNS TO INITIAL YEARS OF FORMAL EDUCATION: HOW BIRTHDATE AFFECTS LATER EDUCATIONAL OUTCOMES

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Economics

at the University of Canterbury

by: Asaad Ismail Ali

Supervised by: Associate Professor Andrea Menclova

Associate Supervision by: Associate Professor Jeremy Clark

University of Canterbury

2019

ACKNOWLEDGEMENTS

First and foremost, I express my deepest and sincere gratitude to my supervisor, Associate Professor Andrea Menclova, for her priceless guidance and mentorship throughout this work. Her understanding and continuous personal guidance has provided a good basis for the present work. She was always there for me from the selection of topic to the completion of the thesis. Her thoughtful and wise comments really made this study a good learning experience for me. I also want to express my gratitude to my associate supervisor Associate Professor Jeremy Clark for his valuable input, guidance and support during this study.

I would like to thank the University of Canterbury, especially the Economics and Finance Department for providing me with the tuition fee waiver and several teaching and research assistantship roles over the past three years. The experience I have gained from working in the department has been influential.

I wish to thank Statistics New Zealand (Stats NZ) for providing me access to use/work on the Integrated Data Infrastructure (IDI) and helping me get acquainted with the IDI lab structure, as well as the New Zealand Association of Economists (NZAE) committee for allowing me to participate in the 2017, 2018, and 2019 conferences and learn from a world class economic research exposure.

Last but not least, I would like to acknowledge my ever praying and encouraging family who was my driving force. Without their continuous prayers, love and encouragement, I wouldn’t have been able to complete this thesis. I would especially like to thank my father who always endowed me with his best wishes and moral support, my mother for always teaching me to do the right thing (no matter the consequences) and my brother for his never-ending care, for which my mere expression of thanks does not suffice.

ABSTRACT

The primary school entry policy in New Zealand is different from the policies of other developed countries like most of the European Union, the United States of America (USA), the United Kingdom, and Australia. In most countries, children start school at a fixed date in contrast to New Zealand where there are rolling admissions and children can start school right after their 5th birthday. Schooling from ages 6 to 16 is compulsory for every child; primary school term 1 in New Zealand starts in February; and the primary education system runs from Years 1 to 8 (Ministry of Education 2015). If a child's birth date is between January and May, the young student will typically spend the year he/she turns 5 in Year 1 and the next year in Year 2. If a child's birth date is between June and December, the student will spend the year he/she turns 5 in Year 0 and start Year 1 the following February. This means that the date of birth of the child affects the number of months/years spent in primary school and may further result in different educational outcomes.

My thesis uses a three-pronged approach to exploit the unusual features of the New Zealand system to test whether additional time spent in school raises subsequent achievement. First, I replicate a Dutch study by Leuven et al. (2010) where a similar system is in operation. In the Netherlands like in New Zealand, schools have a rolling admissions policy and children can start school right after their 4th birthday instead of 5th as in New Zealand. Dutch children with birthdays during, before and after the summer holidays are placed in the same class. Leuven et al. (2010) indicated that these two features of the Dutch schooling system create adequate exogenous variation in children's enrolment opportunities to identify the effects of additional early formal education on later test scores. My thesis replicates Leuven et al. and finds some notable differences. This replication, in addition to being of interest in itself, serves as a useful starting point for analysing the New Zealand school system as the Dutch system is very similar.

Second, I apply a similar approach to New Zealand data and analyse the effects of school start on long-term educational achievement in New Zealand. Specifically, I focus on National Certificate of Educational Achievement (NCEA) and University Entrance (UE) results. Controlling for demographic and socio-economic characteristics, I find that an additional month of schooling (before the start of Year 2) increases the probability of achieving NCEA level 1 by 2.2%, NCEA level 2 by 4.2%, NCEA level 3 by 6.2%, and UE by 5.2%. Thus,

differences in the timing of birth – and hence in school start – seem to have large effects on achievement even years later, in high school.

Finally, I subject my main results to a series of robustness and falsification checks. I also investigate whether the effects are homogeneous across socio-demographic groups or whether they are concentrated in certain sub-populations. I find that the effects are the strongest among male, Māori, and decile 5-7 students.

PUBLICATION AND ACKNOWLEDGMENTS FOR PARTS OF THE THESIS

WORKING PAPER PUBLICATION

- Ali, A. I., & Menclova, A. K. (2018). *Returns to Early Formal Education: A Replication of Leuven et al. (2010)*. Working Papers in Economics 18/19, University of Canterbury, Department of Economics and Finance.

CONFERENCE PRESENTATIONS

- New Zealand Association of Economists Conference (2017): Poster presentation of Chapter 2
- New Zealand Association of Economists Conference (2018): Oral presentation of Chapter 3
- New Zealand Association of Economists Conference (2019): Oral presentation of Chapters 3 and 4

SEMINARS

- Oral presentation of Chapter 2 at the annual PhD Workshop 2017 in Wellington.
- Oral presentation of Chapter 3 in the Department of Economics and Finance, University of Canterbury (June, 2018)

AWARDS

- Statistics New Zealand (2019) prize for "the best use of official statistics" at the New Zealand Association of Economists Conference 2019 for a paper based on Chapters 3 and 4.

STATISTICS NEW ZEALAND DISCLAIMER

The results in this paper are not official statistics. They have been created for research purposes from the IDI, managed by Stats NZ.

The opinions, findings, recommendations, and conclusions expressed in this thesis are those of the author, not Stats NZ, IDI or Ministry of Education (MOE).

Access to the anonymised data used in this study was provided by Statistics NZ under the security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this paper have been confidentialised to protect these groups from identification and to keep their data safe.

Careful consideration has been given to the privacy, security, and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the IDI available from www.stats.govt.nz

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DANS	Data Archiving and Network Services
GDP	Gross Domestic Product
IDI	Integrated Data Infrastructure
IFP	International Fee Paying
LPM	Linear Probability Model
MOE	Ministry of Education
NCEA	National Certificate of Achievement
NSN	National Student Number
NZQA	New Zealand Qualifications Authority
NHL	North American National Hockey League
OECD	Organization for Economic Co-operation and Development
RAE	Relative Age Effect
Stats NZ	Statistics New Zealand
USA	United States of America
UE	University Entrance

CHAPTER 1. INTRODUCTION

Education has a strong correlation with the economic development of any country (Earle, 2010). In modern times when the ‘knowledge economy’ has become a central focus, the role of education becomes all the more prominent in raising human capital. Development agendas rate ‘access to education’ as one of the top priorities (Hanushek & Wößmann, 2007). Furthermore, the quality of education, in addition to simple ‘access’, affects the economic development of any country (Hanushek & Wößmann, 2007).

Early childhood education in particular plays a vital role in developing both the cognitive and the non-cognitive skills of a child (Kautz, Heckman, & Diris, 2014). This is the reason that primary education policies have recently become the main focus in developed countries. James J. Heckman – a Nobel Laureate in Economics – developed a model for the Joint Economic Committee which showed that the formation of life cycle skills is by nature dynamic (Cunha, Heckman, Lochner, & Masterov, 2006). Skill within itself creates skill and similarly motivation within itself causes motivation. If an individual does not become motivated and inspired early on in life to learn and to participate, it is more likely that he/she will fail in economic and social life as an adult.

This thesis focusses on the effects of early formal education on later academic achievement. I use data from the Netherlands (Chapter 2) and New Zealand (Chapters 3 and 4) which share an interesting school start policy; namely, rolling admissions into primary school as soon as a child reaches a certain age. This feature is described in detail below.

New Zealand has high education expenditure to Gross Domestic Product (GDP) ratio when compared to other developed countries. Public spending on education as a proportion of GDP has increased from 5.3% in 2012 to 7.4% in 2016 (UNESCO Institute for Statistics, 2012). In the 2014/15 financial years, core Crown education spending was $12.9 billion (The Treasury - Kaitohutohu Kaupapa Rawa, 2015).

The Netherlands also has a high government expenditure on education. In 2013, its expenditure on education as a percentage of GDP was 5.6% (UNESCO Institute for Statistics, 2016).

In both New Zealand and the Netherlands, students perform well on standardized tests – mathematics, science literacy, and reading tests – compared to other Organization for Economic Co-operation and Development (OECD) countries (Programme for International Student Assessment, 2016). For example, in the Netherlands, 15 year old students score an average of 509 points on a science literacy test compared to the OECD average of 493 points. Students in New Zealand score an average of 513 points in the science literacy test. In mathematics, 15 year old Dutch students score 512 points on average while 15 year old Kiwi students score 495 points on average, as compared to an average of 490 points across OECD countries. Finally, the average reading score of 15 year old Dutch students is 503 points and 15 year old Kiwi students is 509 points compared to 493 points across OECD countries.

The primary school policies of different countries vary in the developed world. Unlike many other developed countries such as the USA, United Kingdom, most of the European Union, and Australia, where schooling starts for all children at a specific date, New Zealand and Dutch schooling officially starts when a child reaches the age of five and four, respectively.

In New Zealand, schooling from ages 6 to 16 is compulsory for every child; primary school term 1 in New Zealand starts in February; and the primary education system runs from Year 1 to Year 8 (Ministry of Education, 2015). If a child's birth date is between January and May, the young student will typically spend the year he/she turns 5 in Year 1 and the next year in Year 2 of primary school. If a child's date of birth is between June and December, the student will usually spend the year he/she turns 5 in Year 0 and start Year 1 the following February. This means that the date of birth of the child affects the number of months/years spent in primary school and may further result in different educational outcomes.

In the Netherlands, the duration of compulsory education is 12 years – from the age of 5 to 16. The academic year for primary to post-secondary education starts in September and ends in June.

This study seeks to make use of the unusual school starting policies of the Netherlands and New Zealand to address a highly contested issue in education: To what extent does early formal education raise human capital, or do schools mainly merely enable high ability students to credibly signal their pre-existing skills? The study is divided into five main chapters. Chapter 1 is an introduction. Chapter 2 consists of a replication of a paper by Leuven et al. (2010). Leuven et al. utilize two features of the Dutch schooling system (based on rolling admissions) that create adequate exogenous variation in children's enrolment opportunities to identify the effects of early formal education on later test scores, measured around the age of 6. Chapter 3 consists of identifying similar effects of school start on long-term educational achievement in New Zealand. Specifically, the focus is on NCEA and UE results, tested around the age of 15-18. Chapter 4 extends the results of Chapter 3 by subjecting them to a set of robustness and falsification checks and investigating homogeneity across socio-demographic groups and sub-populations. Finally, Chapter 5 concludes the thesis.

1.1 RESEARCH GAP

A previous literature in education has shown that in most developed countries (but largely ignoring New Zealand and the Netherlands), primary school entry policies affect later educational outcomes of students with different dates of birth. However, most developed countries start schools at a fixed date in contrast to New Zealand and the Netherlands where school starts when a child turns five and four years of age, respectively. With the different school start policies in mind; this study shows that date of birth has a qualitatively similar impact on later academic achievement in New Zealand and the Netherlands. Importantly, in doing so, it quantifies the (causal) returns to early formal schooling.

1.2 OBJECTIVES

The main objectives of the study are to:

1. Use exogenous variation in children's enrolment opportunities to identify the short-term effects of early formal education on academic achievement (Chapter 2 using data from the Netherlands);
2. Using the same methods, investigate the impacts of early years spent in primary school on longer-term educational outcomes (Chapters 3 and 4 using data from New Zealand).

CHAPTER 2. REPLICATION OF LEUVEN ET AL. (2010)

2.1 INTRODUCTION

A child's rudimentary functions of emotional and cognitive control, communication, and learning are established in the age range from birth to about 6 years – a period also known as early childhood (Heckman, 2000). Psychological adjustment as well as educational achievement are long-term consequences of early childhood development. In recognition of this fact, most developed countries focus on providing free early childhood education (Leuven, Lindahl, Oosterbeek, & Webbink, 2010). This is also the case in the Netherlands where children are able to start attending a publically funded school when they reach the age of four. Although schooling is compulsory from the age of five, 98% of all children start school from the age of four (Leeuwen, Thijs, & Zandbergen, 2009). The school curriculum contains well-thought-out learning activities; by the age of six, children will have typically started to write and read.

The main objective of this study is to replicate and further analyse the paper by Leuven, et al. (2010). Leuven et al. estimated the (short-term) impact of school enrolment opportunities at the age of four on subsequent achievement in language and arithmetic tests in grade 2 (i.e., at the age of six). They exploit two distinct features of the Dutch schooling system that produce exogenous variation in the amount of time children spend in school to identify effects. First, the Netherlands has a rolling admissions policy – children can start school right after their fourth birthday instead of a fixed term start date. Second, children with dates of birth during, before and after the summer holidays are enrolled in the same grade. These two features can create a difference in the time spent in school of up to 11 weeks – and this difference is not a linear function of age. By the time a child is in grade 2 (at the age of six), 11 weeks amount to approximately 15% of his/her schooling time.

Leuven et al. (2010) hypothesise that parents do not take the school admission policy into account when planning a pregnancy/birth. They find support for this hypothesis in that family characteristics are not systematically related to children's birth dates. Treating birth date as exogenous, they then show that a one month increase in the time spent in school results in an increase of 6% of the standard deviation in language scores and 5% of the standard deviation in arithmetic scores, for disadvantaged children[1]. The authors find no effect of additional

1 Edwards (1974) defines disadvantaged children as those children who do not have a home background which prepares them for primary school education. For disadvantaged children, the first few months of primary school may represent a more abrupt change from the standards and values they

schooling caused by birthdate differences for non-disadvantaged children. My replication finds stronger effects of birth date differences than the original study. In particular, I find that, among disadvantaged children, one month of additional schooling increases language test scores by 3% of a standard deviation and arithmetic test scores by 4% of a standard deviation. Among non-disadvantaged children, one month of additional schooling results in an increase of language test scores by 4% of a standard deviation and of arithmetic test scores by 5% of a standard deviation. Hence, I find a large (and statistically significant) effect for both disadvantaged and non-disadvantaged students.

2.2 CONCEPTUAL FRAMEWORK

Primary schooling is compulsory in the Netherlands for every child. Each child can start school when he/she reaches the age of four. Schooling becomes compulsory when the child reaches the age of five. Approximately 98% of students tend to start school at the age of four. Children typically attend primary school from the age of four to twelve; i.e., eight years. When starting school, children are placed in grade one also known as group 1. Every year after the school holidays, children move up one grade. A national examination is conducted at the completion of primary school when the students are twelve years of age. This examination determines which secondary school the student can enrol in.

The exact time and date of school enrolment between the ages of four and five is depend on parent's choice (in addition to the timing of school holidays - which is an important feature for the purposes of this thesis). Another key characteristic of the schooling system is the school year class cohort which consists of children who have birthdays between the 1st of October of a particular year and the 30th of September of the subsequent year. Additionally, each school year is from one summer holiday to the subsequent summer holiday. Every child who joins school after his/her 4th birthday, regardless of the date of joining, spends the first/joining year (till the 30th of September) in grade one. The child then stays in grade one till the next summer holidays. He/she then moves to grade two after the summer holidays. The amount of time each child can spend in school (maximum length of schooling), which is not a linear function of his/her age, differs because of these characteristics.

have grown up with. In Leuven et al. (2010), children are classified as disadvantaged if both parents have at most a degree from a low-level vocational school.

Figure 2.1 (original on the right hand side, replication on the left hand side) shows the association between a child's birth date and his/her potential maximum amount of time spent in school. The segments with slope = 0 are caused by being born during a school holiday and the segments with slope < 0 are caused by being born outside of school holidays. There are five different segments with slope < 0 and five segments with slope = 0, reflecting the timing of school holidays in the Netherlands. Those children who lie on the same 'slope < 0' segment have a one to one relationship between age and the maximum amount of time spent in school; i.e., one extra day of age leads to one extra day spent in school. For these children, any differences in their test scores are attributable to differences in their age as well as differences in their 'maximum length of schooling' (or randomly distributed differences in child/parental/regional characteristics). In contrast, those children who lie on the same 'slope = 0' segment start school on the same day after the holidays. Hence, any systematic differences in test scores among children born in the same segment are attributable to differences in their age only.

2.3 DATA

The data used by Leuven, et al. (2010) is from five waves of a Dutch survey known as the PRIMA Survey which was conducted under the supervision of Data Archiving and Network Services (DANS). It is a time series data set from 1994 to 2004 of each student enrolled in grades 2, 4, 6, and 8 (when standardised tests are taken) in different schools of different provinces/regions in the Netherlands. Each wave consists of approximately 55,000 children in 600 different primary schools. The original study estimates the short term effect of only those students who are in grade 2 and are not repeating the grade. The omission of data on other grades was because of the frequent grade repeaters. Grade repetition was not an issue for children advancing from grade 1 to grade 2. This is because, in the survey, less than 3 percent of the students in second grade are older than what second-graders should be, signalling that they repeated a grade. However, the survey contains no information that can identify whether they repeated the first or the second grade. Therefore, 3 percent is an upper bound on grade repetition from grade 1 to grade 2.For replication purposes, I also analyse children in grade 2 only.

The process of getting access to data used in Leuven, et al. (2010) was very time consuming. Because of the confidentiality of the data, the authors were not allowed to forward the data to other researchers, even for replication purposes. Therefore, I had to apply for data access directly to DANS. Moreover, the data files and the codebooks were all in Dutch so I had to translate them. Finally, the survey contained many variables not required for my replication so I had to filter out and clean the data as well.

Despite my best efforts, there remains a difference in the total number of observations used in the replication and in the original study. The original study has a sample size of 52,835 observations with 28,984 non-disadvantaged children and 23,893 disadvantaged children (11,149 Dutch and 12,744 minority). My replication has only 40,500 observations (12,335 fewer than the original study) - 21,203 for non-disadvantaged children, 8,271 for Dutch disadvantaged and 11,026 for minority disadvantaged. However, my data seem to be largely a random subsample of the data in Leuven et al. as suggested in Tables 2.2 and 2.3.

2.3.1 DESCRIPTION OF VARIABLES

The variables used in the estimation are described in Table 2.1. Like Leuven, et al. (2010), I use *potential* enrolment rather than actual enrolment in school as information about actual enrolment was very limited in the PRIMA survey. More importantly, actual enrolment is very likely to suffer from endogeneity due to parents' choice in timing the start of school of their child. If actual enrolment were available, potential enrolment would need to be used as an instrumental variable. In its absence, Leuven, et al. (2010) essentially use an intent-to-treat approach.

2.4 METHODS & RESULTS

Table 2.2 (original study) and Table 2.3 (replication) show descriptively the characteristics of non-disadvantaged children (column 1), Dutch disadvantaged children (column 2), and minority disadvantaged children (column 3). The main difference between Table 2.2 and Table 2.3 is the number of observations for each category.

The authors of the original study used two tests as outcome measures in their analysis: a language test (which measures understanding of words and sentences, and placement of those words in sentences) and an arithmetic test (which tests classification, counting and ordering). In order to examine children's ability to read, write and master basic math concepts, the Dutch Central Institute for Test Development, in Dutch "Centraal Instituut voor Testontwikkeling" created these tests for the Dutch government. These tests are administered midway through the school year (around February). To compare the achievement levels of different children over the period of time, Leuven, et al. (2010) transform the raw scores of the tests into a standardized measure with standard deviation one and mean zero in each wave. I use the same test outcome measures in this replication.

The exact date of birth is necessary for the analysis but unfortunately was not available in the data set provided by DANS – only the month and year of birth were. Based on descriptive statistics reported in the original Table 1, I mimic Leuven, et al. (2010) by imputing the exact date of birth in two different ways: first, by randomly generating the exact date of birth for each student born in a specific month and, second, by using the midpoint of each month. For instance, if a child's birthday is in May, the mid-point value for his/her date of birth will be May 15th whereas the random value will be a random number between the 1st and the 31st. Table 2.3 shows the descriptive statistics for both these approaches. Since the mean values are very close, I decided to continue my analysis with the randomized date of birth.

Despite the difference in the sample size in each category, I am able to get very similar descriptive statistics to the original study. As shown in Table 2.2 and Table 2.3, the difference in each category is typically not more than a few decimal points.

The average language and arithmetic test scores for different groups are shown in Table 2.2 and Table 2.3 among other descriptive statistics. According to the original study, non-disadvantaged children score higher than the average test score by around one third of a standard deviation. When compared with disadvantaged minority, the non-disadvantaged children score 1.00 standard deviation higher on the language test and 0.80 standard deviation higher on the arithmetic test. When compared with disadvantaged Dutch, the non-disadvantaged children score 0.40 standard deviation higher in both the language and the arithmetic test.

My replication similarly finds that non-disadvantaged children score 0.96 standard deviations higher than the disadvantaged minority children on the language test scores and 0.75 standard deviation higher on the arithmetic test score. On the other hand, when compared with the disadvantaged Dutch, the non-disadvantaged children score 0.35 standard deviations higher on the language test score and 0.42 standard deviations higher on the arithmetic test score.

2.4.1 EXOGENEITY OF THE MAXIMUM LENGTH OF SCHOOLING

Crucial to Leuven et al.'s analysis is their assumption that children's date of birth is exogenous to schooling start dates. To test whether the length of schooling is indeed exogenous, Leuven, et al. (2010) regress the maximum length of schooling on age and age squared, four dummies for each of mothers' and fathers' education, gender dummy, and dummies for regions and years and their interactions. The variation in the length of schooling here comes from the school holidays as shown in Figure 2.1. The results are shown in Table 2.4.

Leuven, et al.'s results show no systematic relationship between background characteristics and the maximum length of schooling variable. This supports the assumption that the maximum length of schooling does not depend on background characteristics and therefore does not affect test scores via these characteristics (rather than directly). This is an intuitive finding as the only way parents could affect the maximum length of schooling would be by timing their birth (and hence the 4th birthdate of their child) with the academic calendar in mind.

I replicate Leuven's Table 2.4 in four different ways. First, in Table 2.5, I conduct an exact replication by regressing the maximum length of schooling on the same variables used by the original study. Second, in Table 2.6, I use mostly the same variables but instead of using region and year-region interaction dummies, I use province dummies and year-province interaction dummies. There are 3 regions in the Netherlands, which are further divided into 12 provinces. Third, I stratify the analysis by gender in Table 2.A.1 – Table 2.A.4 in the Appendix. Finally, in Appendix Table 2.A.5, I change the reference category for mother's education as well as father's education from 'missing' to primary education.

My replication of the date of birth exogeneity checks shows slightly different results. One robust difference is the significance of the gender dummy. Surprisingly, the results show that being a girl increases the potential amount of time spent in school for non-disadvantaged as well as disadvantaged minority children. However, the effect is small. For example, non-disadvantaged girls potentially spend 0.17 more months – or 5 more days - in school than corresponding boys. The results similarly show that mother's upper secondary and higher education as well as father's lower secondary education variables are statistically significant, even if the impacts are quantitatively small. I employ alternative model specifications below in an effort to eliminate any systematic 'sorting'.

When provinces are used instead of regions in Table 2.6, the results show that the gender dummy is still significant with a positive coefficient indicating that girls again have a slightly longer potential time in school than boys. The effect remains small – 5 days at a maximum.

Mother's education and father's education variables also remain statistically significant with quantitatively small effect when provinces are controlled for.

Splitting the sample by gender in Table 2.A.1 – Table 2.A.4 in the Appendix causes other coefficients to become statistically significant and so is downplayed in subsequent analyses. Changing the reference category for mother and father's education from 'missing' to primary education in Table 2.A.5 has little impact on the results.

I move next to Leuven et al.'s key analysis on the effect of time spent in school on subsequent outcomes. The effects of the maximum length of schooling on children's language and arithmetic scores as calculated by Leuven et al. (2010) are shown in Table 2.7 and Table 2.8, respectively. My replications follow in Table 2.9 and Table 2.10. The results are calculated separately for the different categories of children in two different ways: age is entered linearly or also with a squared term on the right-hand side. All regressions for Table 2.7 and Table 2.8 include two region dummies, four year dummies, year-region interactions, a gender dummy, and four dummies each for mother's and father's education. In addition, columns 7 and 8[2] for all of the following tables include a dummy for disadvantaged Dutch as well. The standard errors of all the regression models are robust to heteroscedasticity and corrected for clustering at the school level.

In Leuven et al.'s analysis of language scores in Table 2.7, the results for the linear specification of age and age squared are similar. Therefore, Leuven et al. (2010) conclude that the linear specification of the effect of age is accurate for the given sample. The results show that an additional one month of the maximum length of schooling results in a 5% of a standard deviation increase in the language test score for disadvantaged Dutch children. The corresponding effect for disadvantaged minority children is 7% of a standard deviation. Non-disadvantaged children do not benefit in their language test scores from increases in the maximum length of schooling.

Just as for language scores, the results for arithmetic scores in Table 2.8 for the linear and quadratic age specifications are similar. The results show that an increase in potential time spent in school does not affect arithmetic scores for non-disadvantaged children. On the other hand, a one month increase in maximum schooling for the disadvantaged Dutch children

[2] Columns 7 and 8 are pooled disadvantaged groups

increases their arithmetic score by 7% and for the disadvantaged minority children by 3% of a standard deviation.

My replications are reported in Table 2.9 and Table 2.10. Most notably, I find substantially larger effects than Leuven et al. (2010) for non-disadvantaged children. For disadvantaged children, however, my effects are smaller. For clarity and reassessment of the results, I modify the specifications in Table 2.9 and Table 2.10 in three different ways. First, I run separate models by gender using the same background controls (Table 2.A.6, Table 2.A.7, Table 2.A.11, and Table 2.A.12 in the Appendix). Second, I use 11 province dummies and year-province interactions instead of 2 region dummies and year-region interactions (Table 2.A.8 and Table 2.A.13 in the Appendix). Finally, I estimate separate models by gender using 11 province dummies and year-province interactions instead of region and year-region interactions (Table 2.A.9, Table 2.A.10, Table 2.A.14, and Table 2.A.15 in the Appendix). All of the above specifications produce fairly similar results. Therefore, I focus on the exact replications in my discussion below.

My results for language scores in Table 2.9 show that there is a significant positive impact of the maximum length of schooling for all of the three groups or children. Specifically, the results of my replication indicate that for non-disadvantaged children, one month of additional schooling results in 4% of a standard deviation increase in language test scores. For the disadvantaged Dutch, the effect is nearly identical. For disadvantaged minority children, the effect is slightly weaker: one additional month of schooling results in a 3% of a standard deviation increase in language test scores.

Similarly, for the arithmetic test (Table 2.10), the effects of additional schooling are positive and highly significant for all groups of children. A one month increase in the length of schooling results in a 5% of a standard deviation increase in arithmetic scores for non-disadvantaged children, a 4% increase for the disadvantaged Dutch, and a 3% increase for disadvantaged minority children. Comparing our findings to those in Leuven et al. (2010), some differences could be expected given the different samples and our imprecise measure of the date of birth. Leuven et al. had access to a superior dataset. However, to the extent that our data are largely a random subsample of the data in Leuven at el. (as Tables 2.2 and 2.3 suggest), we would expect to find less precisely estimated effects, not more. Similarly, measurement error in my key right-hand-side variable would bias any estimated effects downwards, not upwards as we find for non- disadvantaged children. Finally, if our sample does differ from Leuven et al. in a systematic way - a possibility which remains, it is unclear why non-disadvantaged children should be so much more affected. Comparing our sample to the original one in Table 2.2, small differences in cohort characteristics appear both in the non-disadvantaged and the disadvantaged group (and, if anything, are larger among disadvantaged children). A future study of the role of early formal education among non-disadvantaged children could reconcile our results.

2.5 CONCLUSIONS

Due to the unique schooling system of the Netherlands, in which children start school on or right after their 4th birthday, Leuven, et al. (2010) were able to introduce an innovative technique to estimate the short-term effect of the potential length of schooling on language and arithmetic test scores, independent from the effect of age. Leuven et al. find that increasing the maximum length of schooling by one month increases language and arithmetic test scores for disadvantaged children by 6% of a standard deviation and 5% of a standard deviation, respectively. They did not find any effects of longer potential schooling for non-disadvantaged children, and concluded that early formal education and the home environment are close substitutes in creating successful performance for non-disadvantaged children. On the other hand, additional schooling environment is better than additional home environment for disadvantaged children.

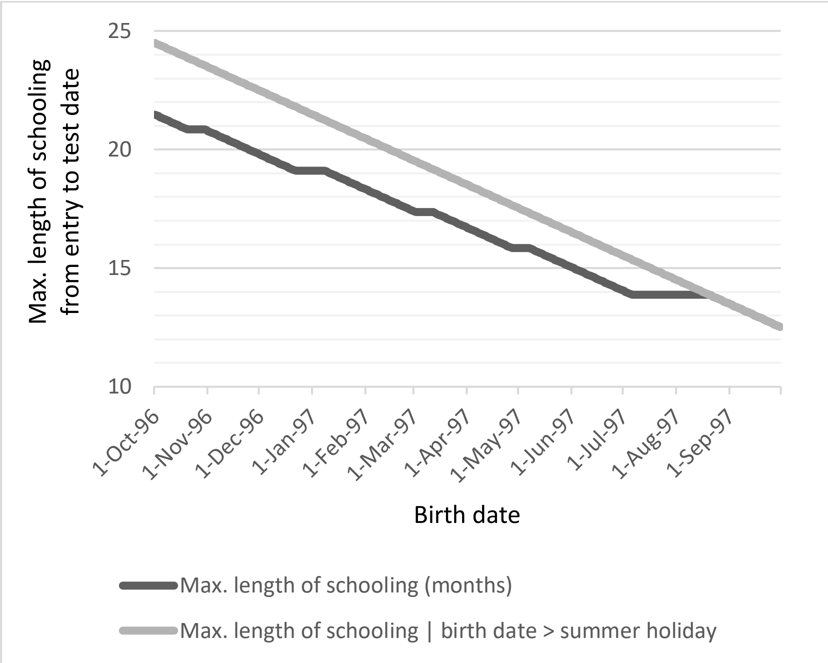
Given that the success gap between non-disadvantaged and disadvantaged children taking language and arithmetic tests in grade 2 is around 60-70% of a standard deviation, an additional month of schooling closes this gap by approximately 10 percent.

My analysis in general endorses the findings of Leuven et al. (2010) for disadvantaged children but I find somewhat smaller effects. Contrary to Leuven et al. (2010), I find that increasing the length of schooling for non-disadvantaged children also results in increasing language and arithmetic test scores and the effect is even larger than for disadvantaged students. In particular, among disadvantaged children, one month of additional schooling increases language test scores by 3% of a standard deviation and arithmetic test scores by 4% of a standard deviation. Among non-disadvantaged children, I find that one month of additional schooling results in an increase of language test scores by 4% of a standard deviation and of arithmetic test scores by 5% of a standard deviation. Hence, in my study, the home environment and the school environment seem to complement each other rather than be substitutes. Among non-disadvantaged students, both environments likely enhance performance. Disadvantaged students also benefit from being in the school environment but to a somewhat smaller degree – possibly due to lower parental effort or availability for complementing school efforts at home.

An increase of around 5-6% of a standard deviation in these test scores comes at a cost of approximately 354 to 541 Euros per child (Leuven, Lindahl, Oosterbeek, & Webbink, 2010). Comparing this to estimates from other countries, Currie and Thomas (1995) estimate the costs of a child participating in the U.S Head Start programme at $3,500 and argue that Head Start participation increases early test scores of disadvantaged white children by 20% of a standard deviation. Therefore, (Leuven, Lindahl, Oosterbeek, & Webbink, 2010) suggest – and I endorse – that instead of targeted programmes such as Head Start, an increase in schooling opportunities for all young children may be a viable and cost-effective alternative.

Figure 2.1: Relationship between the maximum length of schooling and birth date for a cohort

Replication



Leuven et al. (2010)

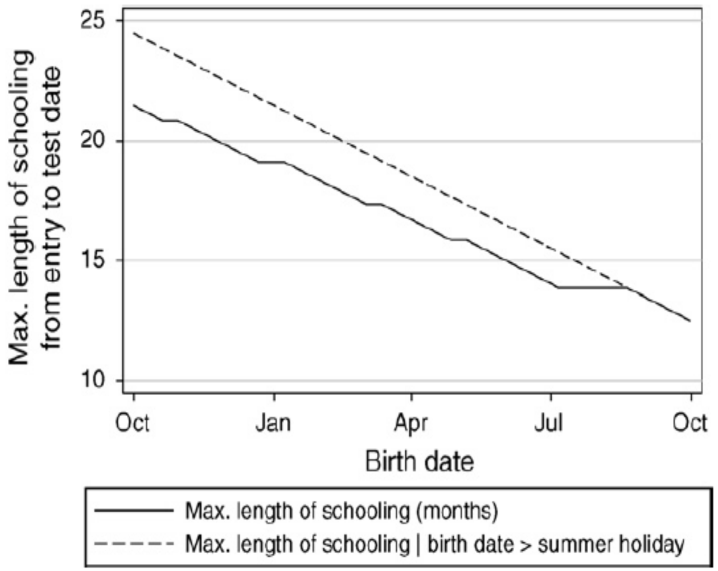


Fig. 1. Relation between birth date and max. length of schooling for a given cohort.

Table 2.1: Variable Descriptions

Name of Variable	**Details**
Student ID	Unique ID for each student
Year	Year (1994-2004)
School NR	Unique number for each school
Group	Class/grade in school
Sex	Gender of the student
OVB-weight (weegfact)	Indicator variables for non-disadvantaged, disadvantaged Dutch, disadvantaged minority student
Date of birth	Date of birth of the student
Date of birth: month	Month of birth of the student
Date of birth: year	Year of birth of the student
Education father	Education of the student's father (primary, lower/upper secondary, higher)
Education mother	Education of the student's mother (primary, lower/upper secondary, higher)
Language score	Language score of the student in group (class)
Arithmetic score	Arithmetic score of the student in group (class)
Penrollr	Potential enrolment in months (time spent in school) – random selection of date
Penrollmp	Potential enrolment in months (time spent in school) – mid-point selection of date

Table 2.2: Descriptive Statistics of Leuven et al. (2010); Table 1 in the original

Descriptive Statistics (mean values and standard deviations)				
	Non.	Disadvantaged		
	Disadv (1)	Dutch (2)	Minority (3)	All (4)
Age (months)	70.330	70.550	70.630	70.590
Standard Deviation	3.400	3.400	3.350	3.380
Max. length of schooling (months)	16.670	16.820	16.890	16.860
Standard Deviation	2.580	2.590	2.560	2.570
Education Mother				
Missing	0.090	0.060	0.090	0.070
Primary	0.010	0.120	0.530	0.340
Lower Secondary	0.160	0.730	0.230	0.470
Upper Secondary	0.500	0.080	0.120	0.100
Higher	0.230	0.010	0.030	0.020
Education Father				
Missing	0.110	0.150	0.180	0.160
Primary	0.010	0.100	0.400	0.260
Lower Secondary	0.190	0.700	0.270	0.470
Upper Secondary	0.410	0.040	0.110	0.070
Higher	0.280	<0.01	0.050	0.030
Not disadvantaged	1	0	0	0
Disadv. Dutch	0	1	0	0.470
Disadv. Minority	0	0	1	0.530
Girl	0.490	0.500	0.500	0.500
Boy	0.510	0.500	0.500	0.500
Language	0.320	-0.040	-0.690	-0.390
Standard Deviation	0.950	0.890	0.830	0.920
Arithmetic	0.280	-0.140	-0.520	-0.340
Standard Deviation	0.980	0.920	0.860	0.910
Number of observations	28,942	11,149	12,744	23,893

Table 2.3: Descriptive Statistics of the Replication

Descriptive Statistics (mean values and standard deviations)					
	Non.	Disadvantaged			Pooled (5)
	Disadv (1)	Dutch (2)	Minority (3)	All (4)	
Age (months)	70.793	71.875	72.486	72.224	71.475
Standard Deviation	4.163	4.763	5.228	5.043	4.659
Max. length of schooling (Penroll)					
random	16.631	16.583	16.616	16.602	16.617
Standard Deviation	2.626	2.611	2.585	2.600	2.612
mid-point	16.656	16.607	16.636	16.624	16.641
Standard Deviation	2.608	2.592	2.564	2.576	2.593
Mother's Education					
Missing	0.090	0.056	0.077	0.068	0.080
Primary	0.010	0.123	0.542	0.362	0.178
Lower Secondary	0.162	0.747	0.236	0.455	0.302
Upper Secondary	0.509	0.067	0.117	0.096	0.312
Higher	0.229	0.008	0.028	0.020	0.129
Father's Education					
Missing	0.112	0.166	0.173	0.170	0.140
Primary	0.010	0.104	0.396	0.271	0.135
Lower Secondary	0.194	0.691	0.275	0.453	0.317
Upper Secondary	0.410	0.035	0.108	0.076	0.251
Higher	0.274	0.004	0.048	0.029	0.157
Non. Disadv	1	0	0	0	0.524
Disadv. Dutch	0	1	0	0.429	0.204
Disadv. Minority	0	0	1	0.571	0.272
Girl	0.484	0.483	0.490	0.487	0.485
Boy	0.517	0.517	0.510	0.513	0.515
Language	0.347	-0.007	-0.609	-0.351	0.014
Standard Deviation	0.949	0.886	0.834	0.907	0.992
Arithmetic	0.306	-0.112	-0.441	-0.300	0.018
Standard Deviation	0.989	0.905	0.856	0.892	0.991
Number of observations	21,203	8,271	11,026	19,297	40,500

Table 2.4: Exogeneity Check in Leuven et al. (2010); Table 2 in the original

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	0.005	-0.002	-0.005	-0.005	-0.002
Standard Errors	0.014	0.011	0.007	0.006	0.005
Lower Secondary	0.006	-0.003	0.007	0.002	0.004
Standard Errors	0.007	0.009	0.008	0.006	0.004
Upper Secondary	0.006	0.000	0.002	0.002	0.004
Standard Errors	0.007	0.011	0.009	0.007	0.004
Higher	0.011	0.028	0.005	0.010	0.009*
Standard Errors	0.007	0.022	0.014	0.011	0.005
Education Father (reference category - missing)					
Primary	-0.010	-0.004	0.006	0.003	0.002
Standard Errors	0.013	0.009	0.006	0.005	0.004
Lower Secondary	-0.006	-0.004	-0.003	-0.004	-0.004
Standard Errors	0.006	0.006	0.006	0.004	0.004
Upper Secondary	-0.007	-0.007	-0.005	-0.006	-0.005
Standard Errors	0.006	0.011	0.008	0.006	0.004
Higher	-0.007	-0.033	-0.007	-0.011	-0.006
Standard Errors	0.006	0.030	0.011	0.009	0.004
Disadv. Dutch				0.000	-0.003
Standard Errors				0.003	0.003
Disadv. Minority					-0.003
Standard Errors					0.003
Girl	0.003	-0.006	0.004	0.000	0.002
Standard Errors	0.002	0.004	0.003	0.003	-0.002
N	28,942	11,149	12,744	23,893	52,835
F-test joint sign	0.834	0.819	0.490	0.660	0.369

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school.

Table 2.5: Exogeneity Check (exact replication)

Maximum length of schooling and background characteristics					
	Non. Disadv (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	-0.213	0.029	0.086	0.053	-0.052
Standard Errors	0.183	0.125	0.106	0.081	0.066
Lower Secondary	-0.063	0.025	0.164	0.118	0.037
Standard Errors	0.092	0.113	0.109	0.079	0.058
Upper Secondary	-0.020	-0.082	0.259*	0.174	0.076
Standard Errors	0.085	0.150	0.113	0.092	0.057
Higher	0.075	0.245	0.344*	0.332*	0.161*
Standard Errors	0.086	0.315	0.163	0.141	0.062
Education Father (reference category - missing)					
Primary	0.293	-0.200	-0.100	-0.123*	-0.090
Standard Errors	0.191	0.109	0.072	0.062	0.056
Lower Secondary	-0.039	-0.114	-0.138*	-0.132**	-0.105*
Standard Errors	0.079	0.073	0.069	0.050	0.041
Upper Secondary	0.010	0.100	-0.063	-0.028	-0.040
Standard Errors	0.080	0.152	0.093	0.083	0.047
Higher	0.109	0.978*	0.033	0.083	0.053
Standard Errors	0.077	0.491	0.118	0.108	0.049
Disadv. Dutch				-0.014	-0.160***
Standard Errors				0.044	0.039
Disadv. Minority					-0.135**
Standard Errors					0.045
Girl	0.166***	0.073	0.129**	0.106**	0.133***
Standard Errors	0.028	0.047	0.043	0.032	0.022
No. of observations	21,203	8,271	11,026	19,297	40,500

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.6: Exogeneity Check (replication with province dummies)

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	-0.206	0.048	0.079	0.055	-0.046
Standard Errors	0.183	0.129	0.106	0.081	0.065
Lower Secondary	-0.050	0.035	0.158	0.120	0.040
Standard Errors	0.093	0.116	0.109	0.079	0.058
Upper Secondary	-0.008	-0.077	0.252*	0.171	0.080
Standard Errors	0.086	0.153	0.113	0.093	0.057
Higher	0.089	0.251	0.347*	0.332*	0.166**
Standard Errors	0.087	0.323	0.161	0.140	0.062
Education Father (reference category - missing)					
Primary	0.275	-0.185	-0.101	-0.122	-0.094
Standard Errors	0.192	0.108	0.072	0.062	0.056
Lower Secondary	-0.047	-0.107	-0.139*	-0.131**	-0.107*
Standard Errors	0.081	0.073	0.068	0.050	0.041
Upper Secondary	0.002	0.118	-0.064	-0.029	-0.044
Standard Errors	0.081	0.146	0.093	0.084	0.048
Higher	0.099	1.039*	0.035	0.082	0.049
Standard Errors	0.078	0.478	0.119	0.109	0.050
Disadv. Dutch				-0.003	-0.161***
Standard Errors				0.046	0.039
Disadv. Minority					-0.137**
Standard Errors					0.045
Girl	0.165***	0.074	0.129**	0.107**	0.132***
Standard Errors	0.028	0.048	0.043	0.032	0.022
No. of observations	21,203	8,271	11,026	19,297	40,500

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.7: Effect of the maximum length of schooling on language scores in Leuven et al. (2010); Table 3 in the original

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	-0.037*	-0.021	0.047	0.053	0.074***	0.066*	0.061	0.060**
Standard Errors	0.021	0.027	0.032	0.039	0.029	0.034	0.021	0.026
Age	0.086***	0.168**	0.024	0.057	-0.003	-0.048	0.010	0.008
Standard Errors	0.016	0.078	0.024	0.116	0.022	0.102	0.016	0.075
Age squared/100		-0.067		-0.027		0.036		0.002
Standard Errors		0.063		0.092		0.080		0.060
R-squared	0.082	0.082	0.084	0.084	0.090	0.090	0.195	0.195

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.8: Effect of the maximum length of schooling on arithmetic scores in Leuven, et al. (2010); Table 4 in the original

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	-0.027	-0.007	0.069**	0.065	0.052*	0.031	0.060***	0.047*
Standard Errors	0.02	0.025	0.035	0.042	0.029	0.037	0.022	0.027
Age	0.089***	0.192**	0.014	-0.009	0.019	-0.095	0.017	-0.049
Standard Errors	0.015	0.084	0.026	0.122	0.022	0.105	0.017	0.078
Age squared/100		-0.083		0.018		0.092		0.054
Standard Errors		0.067		0.097		0.086		0.063
R-squared	0.093	0.093	0.087	0.087	0.086	0.086	0.123	0.123

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.9: Effect of the maximum length of schooling on language scores (exact replication)

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.048***	0.038***	0.045***	0.034***	0.031***	0.024***	0.037***	0.029***
Standard Errors	0.003	0.003	0.004	0.005	0.003	0.004	0.003	0.003
Age	0.018***	0.462***	0.018***	0.307***	0.019***	0.211***	0.018***	0.242***
Standard Errors	0.002	0.040	0.002	0.058	0.002	0.039	0.001	0.032
Age squared/100		-0.308***		-0.196***		-0.129***		-0.151***
Standard Errors		0.028		0.039		0.026		0.021
R-squared	0.075	0.080	0.072	0.076	0.066	0.069	0.162	0.165

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.10: Effect of the maximum length of schooling on arithmetic scores (exact replication)

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.058***	0.046***	0.053***	0.040***	0.032***	0.024***	0.041***	0.032***
Standard Errors	0.004	0.004	0.004	0.004	0.004	0.004	0.003	0.003
Age	0.021***	0.545***	0.020***	0.380***	0.023***	0.256***	0.022***	0.296***
Standard Errors	0.002	0.046	0.002	0.061	0.002	0.038	0.001	0.031
Age squared/100		-0.363***		-0.244***		-0.157***		-0.185***
Standard Errors		0.032		0.041		0.025		0.021
R-squared	0.078	0.085	0.077	0.082	0.060	0.064	0.091	0.095

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

2.6 APPENDIX

Table 2.A.1: Exogeneity Check (for boys only with region dummies)

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	-0.233	0.237	0.090	0.120	-0.024
Standard Errors	0.231	0.172	0.133	0.100	0.082
Lower Secondary	-0.091	0.288*	0.277*	0.269**	0.133
Standard Errors	0.113	0.141	0.133	0.097	0.070
Upper Secondary	-0.058	0.174	0.306*	0.286*	0.145*
Standard Errors	0.102	0.195	0.148	0.117	0.068
Higher	0.027	0.916*	0.353	0.474*	0.230**
Standard Errors	0.106	0.442	0.231	0.206	0.079
Education Father (reference category - missing)					
Primary	0.493	-0.293*	-0.070	-0.162	-0.090
Standard Errors	0.278	0.144	0.095	0.084	0.077
Lower Secondary	0.016	-0.281**	-0.101	-0.200**	-0.146*
Standard Errors	0.102	0.104	0.093	0.069	0.057
Upper Secondary	0.052	-0.369	-0.087	-0.185	-0.110
Standard Errors	0.099	0.235	0.128	0.118	0.065
Higher	0.176	1.116	0.001	-0.009	0.023
Standard Errors	0.097	0.865	0.147	0.139	0.066
Disadv. Dutch				-0.024	-0.147**
Standard Errors				0.059	0.077
Disadv. Minority					-0.125*
Standard Errors					0.054
No. of observations	10,951	4,274	5,619	9,893	20,844

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.2: Exogeneity Check (for girls only with region dummies)

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	-0.244	-0.254	0.080	-0.019	-0.087
Standard Errors	0.282	0.171	0.140	0.112	0.093
Lower Secondary	-0.060	-0.293	0.044	-0.042	-0.067
Standard Errors	0.139	0.151	0.147	0.107	0.081
Upper Secondary	-0.000	-0.387	0.210	0.055	0.002
Standard Errors	0.128	0.214	0.159	0.132	0.080
Higher	0.101	-0.538	0.325	0.173	0.085
Standard Errors	0.133	0.332	0.214	0.183	0.089
Education Father (reference category - missing)					
Primary	0.121	-0.112	-0.128	-0.084	-0.096
Standard Errors	0.248	0.146	0.107	0.088	0.077
Lower Secondary	-0.084	0.038	-0.180	-0.067	-0.063
Standard Errors	0.122	0.100	0.100	0.071	0.059
Upper Secondary	-0.018	0.531**	-0.036	0.130	0.036
Standard Errors	0.122	0.191	0.130	0.109	0.068
Higher	0.051	1.007	0.073	0.186	0.085
Standard Errors	0.120	0.685	0.170	0.160	0.072
Disadv. Dutch				0.000	-0.174**
Standard Errors				0.060	0.053
Disadv. Minority					-0.146*
Standard Errors					0.063
No. of observations	10,252	3,997	5,407	9,404	19,656

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.3: Exogeneity Check (for boys only with provinces dummies)

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	-0.233	0.237	0.090	0.120	-0.024
Standard Errors	0.231	0.172	0.133	0.100	0.082
Lower Secondary	-0.091	0.288*	0.277*	0.269**	0.133
Standard Errors	0.113	0.141	0.133	0.097	0.070
Upper Secondary	-0.058	0.174	0.306*	0.286*	0.145*
Standard Errors	0.102	0.195	0.148	0.117	0.068
Higher	0.027	0.916*	0.353	0.474*	0.230**
Standard Errors	0.106	0.442	0.231	0.206	0.079
Education Father (reference category - missing)					
Primary	0.493	-0.293*	-0.070	-0.162	-0.090
Standard Errors	0.278	0.144	0.095	0.084	0.077
Lower Secondary	0.016	-0.281**	-0.101	-0.200**	-0.146*
Standard Errors	0.102	0.104	0.093	0.069	0.057
Upper Secondary	0.052	-0.369	-0.087	-0.185	-0.110
Standard Errors	0.099	0.235	0.128	0.118	0.065
Higher	0.176	1.116	0.001	-0.009	0.023
Standard Errors	0.097	0.865	0.147	0.139	0.066
Disadv. Dutch				-0.024	-0.147**
Standard Errors				0.059	0.053
Disadv. Minority					-0.125*
Standard Errors					0.054
No. of observations	10,951	4,274	5,619	9,893	20,844

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.4: Exogeneity Check (for girls only with provinces dummies)

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - missing)					
Primary	-0.244	-0.254	0.080	-0.019	-0.087
Standard Errors	0.282	0.171	0.140	0.112	0.093
Lower Secondary	-0.060	-0.292	0.044	-0.042	-0.067
Standard Errors	0.139	0.151	0.147	0.107	0.081
Upper Secondary	-0.000	-0.387	0.210	0.055	0.002
Standard Errors	0.128	0.214	0.159	0.132	0.080
Higher	0.101	-0.538	0.325	0.173	0.085
Standard Errors	0.133	0.332	0.214	0.183	0.089
Education Father (reference category - missing)					
Primary	0.121	-0.112	-0.128	-0.084	-0.096
Standard Errors	0.248	0.146	0.107	0.088	0.077
Lower Secondary	-0.084	0.038	-0.180	-0.067	-0.063
Standard Errors	0.122	0.100	0.100	0.071	0.059
Upper Secondary	-0.018	0.531**	-0.036	0.130	0.036
Standard Errors	0.122	0.191	0.130	0.109	0.068
Higher	0.051	1.007	0.073	0.186	0.085
Standard Errors	0.120	0.685	0.170	0.160	0.072
Disadv. Dutch				0.000	-0.174**
Standard Errors				0.060	0.053
Disadv. Minority					-0.146*
Standard Errors					0.063
No. of observations	10,252	3,997	5,407	9,404	19,656

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.5: Exogeneity Check (Reference category for mother's and father's education changed)

Maximum length of schooling and background characteristics					
	Non. Disadv. (1)	Disadvantaged			Pooled (5)
		Dutch (2)	Minority (3)	All (4)	
Education Mother (reference category - primary)					
Lower Secondary	0.150	-0.004	0.078	0.065	0.089
Standard Errors	0.167	0.084	0.058	0.048	0.045
Upper Secondary	0.193	-0.111	0.173*	0.121	0.128*
Standard Errors	0.165	0.127	0.077	0.069	0.049
Higher	0.288	0.216	0.258	0.279*	0.213***
Standard Errors	0.165	0.297	0.141	0.124	0.057
Missing	0.213	-0.029	-0.086	-0.053	0.052
Standard Errors	0.183	0.125	0.106	0.081	0.066
Education Father (reference category - primary)					
Lower Secondary	-0.331	0.085	-0.038	-0.009	-0.014
Standard Errors	0.167	0.092	0.065	0.055	0.050
Upper Secondary	-0.283	0.300	0.037	0.094	0.051
Standard Errors	0.175	0.153	0.082	0.074	0.051
Higher	-0.184	1.178*	0.133	0.206*	0.144*
Standard Errors	0.181	0.493	0.109	0.103	0.057
Missing	-0.293	0.199	0.100	0.123	0.090
Standard Errors	0.191	0.109	0.072	0.062	0.056
Disadv. Dutch				-0.014	-0.160***
Standard Errors				0.044	0.039
Disadv. Minority					-0.135**
Standard Errors					0.045
Girl	0.166***	0.073	0.129**	0.106**	0.133***
Standard Errors	0.028	0.047	0.043	0.032	0.022
No. of observations	21,203	8,271	11,026	19,297	40,500

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.6: Effect of maximum length of schooling on language scores for boys with same background characteristics as the original study

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.043***	0.031***	0.039***	0.028***	0.028***	0.019***	0.034***	0.024***
Standard Errors	0.004	0.004	0.005	0.006	0.005	0.005	0.004	0.004
Age	0.019***	0.472***	0.022***	0.345***	0.020***	0.255***	0.021***	0.289***
Standard Errors	0.002	0.054	0.003	0.072	0.002	0.052	0.002	0.042
Age squared/100		-0.312***		-0.218***		-0.157***		-0.181***
Standard Errors		0.037		0.049		0.035		0.028
R-squared	0.059	0.065	0.069	0.073	0.065	0.069	0.153	0.157

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.7: Effect of maximum length of schooling on language scores for girls with same background characteristics as the original study

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.056***	0.048***	0.051***	0.042***	0.033***	0.028***	0.041***	0.035***
Standard Errors	0.004	0.004	0.006	0.007	0.004	0.005	0.004	0.004
Age	0.015***	0.484***	0.013***	0.292**	0.017***	0.168**	0.016***	0.199***
Standard Errors	0.003	0.058	0.003	0.088	0.002	0.054	0.002	0.044
Age squared/100		-0.327***		-0.191***		-0.102***		-0.124***
Standard Errors		0.041		0.059		0.036		0.029
R-squared	0.068	0.073	0.067	0.070	0.070	0.071	0.169	0.170

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.8: Effect of maximum length of schooling on language scores with different background characteristics

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.049***	0.039***	0.045***	0.035***	0.031***	0.025***	0.038***	0.030***
Standard Errors	0.003	0.003	0.004	0.005	0.003	0.004	0.003	0.003
Age	0.018***	0.465***	0.018***	0.301***	0.019***	0.208***	0.018***	0.240***
Standard Errors	0.002	0.040	0.002	0.058	0.002	0.040	0.001	0.031
Age squared/100		-0.310***		-0.192***		-0.127***		-0.150***
Standard Errors		0.028		0.039		0.026		0.021
R-squared	0.082	0.087	0.082	0.086	0.076	0.079	0.169	0.171

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. Different background characteristics means using provinces dummies instead of region dummies.

Table 2.A.9: Effect of maximum length of schooling on language scores with different background characteristics for boys

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.043***	0.031***	0.040***	0.029***	0.028***	0.020***	0.034***	0.025***
Standard Errors	0.004	0.004	0.005	0.006	0.005	0.005	0.004	0.004
Age	0.019***	0.470***	0.022***	0.335***	0.021***	0.251***	0.021***	0.284***
Standard Errors	0.002	0.054	0.003	0.070	0.002	0.051	0.002	0.041
Age squared/100		-0.311***		-0.212***		-0.154***		-0.177***
Standard Errors		0.037		0.048		0.034		0.028
R-squared	0.069	0.075	0.085	0.089	0.075	0.079	0.162	0.165

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. Different background characteristics means using provinces dummies instead of region dummies.

Table 2.A.10: Effect of maximum length of schooling on language scores with different background characteristics for girls

Language								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.057***	0.048***	0.051***	0.042***	0.035***	0.030***	0.041***	0.035***
Standard Errors	0.004	0.004	0.006	0.007	0.004	0.005	0.004	0.004
Age	0.015***	0.494***	0.014***	0.283**	0.017***	0.164**	0.016***	0.196***
Standard Errors	0.003	0.059	0.003	0.088	0.002	0.053	0.002	0.043
Age squared/100		-0.333***		-0.184**		-0.099**		-0.123***
Standard Errors		0.041		0.059		0.036		0.029
R-squared	0.078	0.083	0.079	0.082	0.084	0.086	0.177	0.178

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. Different background characteristics means using provinces dummies instead of region dummies.

Table 2.A.11: Effect of maximum length of schooling on arithmetic scores for boys with same background characteristics as the original study

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.055***	0.041***	0.043***	0.029***	0.029***	0.020***	0.036***	0.025***
Standard Errors	0.005	0.005	0.005	0.006	0.005	0.005	0.004	0.004
Age	0.021***	0.552***	0.027***	0.427***	0.024***	0.282***	0.025***	0.341***
Standard Errors	0.003	0.067	0.003	0.077	0.002	0.052	0.002	0.043
Age squared/100		-0.366***		-0.271***		-0.173***		-0.213***
Standard Errors		0.046		0.052		0.035		0.029
R-squared	0.071	0.079	0.083	0.090	0.063	0.067	0.090	0.096

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. Different background characteristics means using provinces dummies instead of region dummies.

Table 2.A.12: Effect of maximum length of schooling on arithmetic scores for girls with same background characteristics as the original study

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.062***	0.053***	0.065***	0.053***	0.035***	0.028***	0.047***	0.039***
Standard Errors	0.005	0.005	0.006	0.007	0.005	0.005	0.004	0.004
Age	0.020***	0.561***	0.011***	0.367***	0.022***	0.229***	0.018***	0.258***
Standard Errors	0.003	0.059	0.003	-0.242***	0.002	0.046	0.002	0.041
Age squared/100		-0.378***		0.063		-0.140***		-0.163***
Standard Errors		0.041				0.031		0.028
R-squared	0.088	0.095	0.079	0.084	0.062	0.064	0.094	0.097

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.13: Effect of maximum length of schooling on arithmetic scores with different background characteristics

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.059***	0.046***	0.053***	0.041***	0.032***	0.024***	0.041***	0.032***
Standard Errors	0.003	0.004	0.004	0.004	0.004	0.004	0.003	0.003
Age	0.020***	0.550***	0.021***	0.375***	0.023***	0.258***	0.022***	0.298***
Standard Errors	0.002	0.045	0.002	0.060	0.002	0.038	0.001	0.031
Age squared/100		-0.367***		-0.241***		-0.158***		-0.187***
Standard Errors		0.031		0.041		0.025		0.021
R-squared	0.087	0.093	0.097	0.102	0.069	0.073	0.101	0.105

Note: All regressions include 4 year dummies, 2 region dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 2.A.14: Effect of maximum length of schooling on arithmetic scores with different background characteristics for boys

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.055***	0.040***	0.044***	0.030***	0.029***	0.020***	0.036***	0.025***
Standard Errors	0.005	0.005	0.005	0.006	0.005	0.005	0.004	0.004
Age	0.021***	0.555***	0.027***	0.419***	0.024***	0.284***	0.025***	0.341***
Standard Errors	0.002	0.067	0.003	0.076	0.002	0.052	0.002	0.043
Age squared/100		-0.368***		-0.265***		-0.174***		-0.213***
Standard Errors		0.046		0.051		0.035		0.029
R-squared	0.082	0.090	0.109	0.115	0.075	0.079	0.102	0.107

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. Different background characteristics means using provinces dummies instead of region dummies.

Table 2.A.15: Effect of maximum length of schooling on arithmetic scores with different background characteristics for girls

Arithmetic								
	Non. Disadv.		Disadvantaged					
			Dutch		Minority		All	
	1	2	3	4	5	6	7	8
Max. length of schooling	0.063***	0.054***	0.065***	0.054***	0.036***	0.029***	0.047***	0.039***
Standard Errors	0.005	0.005	0.006	0.007	0.005	0.005	0.004	0.004
Age	0.020***	0.568***	0.012***	0.358***	0.021***	0.226***	0.018***	0.255***
Standard Errors	0.003	0.058	0.003	0.093	0.002	0.046	0.002	0.041
Age squared/100		-0.383***		-0.236***		-0.138***		-0.160***
Standard Errors		0.040		0.063		0.031		0.028
R-squared	0.097	0.104	0.105	0.109	0.073	0.076	0.108	0.111

Note: All regressions include 4 year dummies, 11 province dummies and their interactions, age and age squared. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. Different background characteristics means using provinces dummies instead of region dummies.

CHAPTER 3. EFFECTS OF SCHOOL START ON LONG-TERM EDUCATIONAL ACHIEVEMENT IN NEW ZEALAND

3.1 INTRODUCTION

Unlike many other developed countries such as the USA, United Kingdom, most of the European Union, and Australia, where schooling starts for all children on a specific date, New Zealand schooling officially starts when a child reaches the age of five. Schooling from ages 6 to 16 is compulsory for every child. Primary school term 1 in New Zealand starts in February and the primary education system goes from Year 0 to Year 8 (Ministry of Education, 2015). If a child's birth date is between January and May, the young student will typically spend the year he/she turns 5 in Year 1 and the next year in Year 2. If a child's birth date is between June and December, the student will usually spend the year he/she turns 5 in Year 0 and start Year 1 the following February. This means that the date of birth of the child affects the number of months/years spent in primary school and may further result in different educational outcomes.

The main objective of this chapter is to make use of the unusual school start policy to study the effects of early school attendance on the individual's later educational outcomes at the end of high school, measured by NCEA and UE results. The study finds large positive returns to early schooling.

3.2 LITERATURE REVIEW

There are three different aspects of school start that have been examined previously in regards to educational achievement. The first is the effect of *absolute age*; i.e., different children start school at a slightly different age and hence at a different stage of their cognitive and social development (mechanism A). The second is the difference in *relative age* among children starting school; i.e., some are younger/older than their peers (mechanism B). Finally, there is the causal effect of schooling on educational outcomes (mechanism C).

3.2.1 ABSOLUTE AND RELATIVE AGE EFFECTS (MECHANISMS A & B)

A number of studies find better academic achievement among children starting school at an older age. Strøm (2004) uses Norwegian data to explore the relationship between the age at which children's formal schooling begins and children's achievements towards the completion of early schooling - holding the date at which school starts constant. Strøm's study determines that younger students have a considerable disadvantage compared to older peers within the same class. The oldest students, born in January, generally score higher in reading tests at 15 to 16 years of age. Compared to the youngest students, born in December, their scores are higher by around 20% of the standard deviation. Strøm adds that he is unable to propose a substitute school start policy which can eliminate this disadvantage.

Datar (2004) examines the effect of postponing kindergarten admission in the USA on children's academic success. Using instrumental variables based on an exogenous discrepancy in birth dates and kindergarten admission age policies, Datar finds that starting kindergarten a year older considerably improves test scores at kindergarten admission. More importantly, the trajectory of test scores is steeper during the first two years of primary school for older children. Datar also suggests that the advantages of delaying kindergarten admission tend to be considerably higher for at-risk, such as poor and disabled, children.

Kawaguchi (2011) uses a Japanese labour force survey to demonstrate that older students in a school group have superior educational achievement and labour market outcomes compared to their younger peers.

Crawford, Dearden and Greaves (2013) show that the oldest children in a particular academic year in England perform considerably better than the youngest children in national achievement tests until the age of 19. Importantly, this difference is experienced when the students turn 16 and make decisions about continuing further secondary school studies as well as when they turn 19 and make decisions about higher education.

Using USA data, Lubotsky and Kaestner (2016) examine whether children with a high level of cognitive and non-cognitive skills at the start of kindergarten experience higher gains in these skills in subsequent years. They show that older kids in kindergarten score higher than the younger ones on both cognitive and non-cognitive measures of achievement. Their cognitive assessment scores grow quicker during kindergarten and first grade. However, the younger entrants start doing better after the first grade and their scores catch up.

However, the positive effect of older age at school start is not observed universally. For example, Angrist and Krueger (1992) examine the effects of the age at school start on later academic performance in the USA. To get exogenous variation in the age at school start (and hence causal effects), they use mandatory school attendance laws as an instrumental variable. Unlike previous studies (e.g., DiPasquale, Moule,, & Flewelling, 1980; Warren, Levin, & Tyler, 1986) which used children's primary school test scores as the outcome variable, Angrist et al. (1992) argue that a superior measure of academic achievement than aptitude test performance at an early age may be the years of education that a child eventually attains. Their results show that children who enter school older attain relatively less education.

Zhang, Zhong and Zhang (2017) use the China Education Panel Survey to test the effect of school starting age on junior high school academic achievement. The results of their study show that a one-year delay in starting school decreases student's cognitive scores in 7th grade by 0.303 standard deviations. They further investigate the mechanisms underlying the relationship between age at entrance and educational outcomes and find that the decrease in scores depends on the accumulation of human capital prior to the start of primary school. In the absence of preschools in China, wealthier parents invest a lot more in their children's pre-school development as compared to poor parents.

The relative age effect (RAE) is also very common and an inescapable phenomenon in competitive sports. Musch and Grondin (2001) show this by reviewing a wide variety of sports studies on the RAEs. They show that RAEs are a common phenomenon in competitive sports. They suggest that bringing RAEs to the attention of all coaches and team managers in the minor sports system is a necessary first step towards safeguarding equal treatment and unbiased competition among players. Barnsley and Thompson (1988) show RAEs in minor hockey. As younger children are at an earlier stage of development than their larger/stronger team members, they are more likely to experience failure and frustration and hence grow an inferior expectation of themselves as hockey players. Boucher and Mutimer (1994) replicate a series of studies (Barnsley & Thompson, 1988; Barnsley, Thompson, & Barnsley, 1985; Daniel & Janssen, 1987; Grondin, Deshaies, & Nault, 1984) of professional ice-hockey players and, like the original studies, find a strong connection between relative age of the players and their participation and contribution in the sport. Cobley et al. (2009) confirm the presence of RAEs through a meta-analytical review of 38 studies, spanning 1984 to 2007, consisting of 253 independent samples across 14 sports and 16 different countries. Fumarco et al. (2017), on the other hand, find an inverse RAE in the North American National Hockey League (NHL); i.e. players born in the last quarter of a calendar year score more and have higher earnings than those born in the first quarter.

It is clear from the above articles that age may have a significant role to play in sports as well as the educational/academic achievement of students. Hence, it is vital to control for the students' age in my analyses. However, given the constant school start date in most countries, the above articles cannot 1) examine whether gradual admittance into early primary education – at a *constant age* – eliminates the effect of a student's date of birth on later educational attainment or 2) study the causal effect of the time spent in school on later outcomes. I turn to these issues below.

3.2.2 THE EFFECT OF THE LENGTH OF SCHOOLING (MECHANISM C)

There are a few studies that try to estimate the causal effect of time spent in school on educational outcomes. These studies use different identification techniques. Some use data on students of the same age but in different grades, i.e. comparable cognitive skills but a different level of education, while others use a unique school system that allows students to enter school at a certain age instead of a certain date.

Cahan and Cohen (1989) estimate the effects of both age and time spent in school for over 12,000 students in grades four to six in Israel. The effect of age is measured as the difference in mean predicted scores between the youngest and the oldest students in a particular grade whereas the effect of time spent in school is measured as the difference in mean predicted scores between the oldest student in that grade and the youngest student in the higher adjacent grade. The authors conclude that one additional year of schooling increases test scores by 0.30 of a standard deviation. On the other hand, being a year older increases the test scores by 0.15 of a standard deviation. Therefore, the effect of an additional year of schooling is on average about twice the effect of being a year older.

Cliffordson and Gustafsson (2008) estimate the effects of both age and schooling on various aspects of intellectual performance in Sweden. They base their analysis on the test scores from military enlistment measuring ‘General visualization ability’, ‘Crystallized intelligence’ and ‘Fluid ability’ at age 16. The tests occur on different dates throughout the year that gives differences in both age and length of schooling among individuals at the time of the test. The authors find that both schooling and age generally raise performance, with the effect of schooling being considerably higher than the effect of age.

Most relevant for my study, Leuven et al. (2010) evaluate the effect of expanding possibilities for early enrolment at school on early achievement using a novel quasi-experimental strategy. They exploit two distinct features of the Dutch schooling system. One is their rolling admissions policy; i.e. children do not have to wait to start primary school on a particular date, they can start right after their fourth birthday. Second, children with birthdays during or right after school holidays start at the same time (at the beginning of the next term) and are put in the same class. The authors use the exogenous variation created by these distinct features in children’s enrolment opportunities to identify their effects on subsequent test scores. They conclude that an additional month of schooling for disadvantaged children increases their arithmetic test scores by 5 percent of a standard deviation and their language test scores by 6 percent of a standard deviation. The study finds no effects for non-disadvantaged children.

Chapter 2 – and Ali and Menclova (2018) – replicate Leuven's study. This replication in general endorses the findings of Leuven et al. but with some notable differences. Specifically, as discussed in Chapter 2, I find positive effects of the time spent in school for both disadvantaged and non-disadvantaged children. On average, an additional month of schooling for disadvantaged children increases their arithmetic and language test scores by three percent of a standard deviation. An additional month of schooling for non-disadvantaged children increases their arithmetic test scores by five percent of a standard deviation and their language test scores by four percent of a standard deviation.

For completeness, other studies suggest that early school attendance may have long-term effects beyond academic achievement. For example, Lleras-Muney (2005) shows a large casual effect of education on mortality in the USA. The author estimates the effect using two different ways: GLS and IV estimation. The results from the GLS estimation show that the probability of dying in the next ten years decreases by about 1.3 percentage points with an additional year of education. The IV estimation shows a much larger effect: an additional year of schooling decreases the probability of dying in the next 10 years by about 3.6 percentage points. The study further elaborates on how life expectancy gains can arise from this effect. It shows that in 1960, at age 35, an additional year of education increased the life expectancy by as much as 1.7 years.

3.3 IDENTIFICATION STRATEGY

As noted above, in New Zealand, the timing of birth – and hence a child's fifth birthday – affects how much time an individual spends in early primary education. If a child is born between January and May, he/she will typically start school in Year 1 and will move to Year 2 the subsequent year. If a child is born between June and December, he/she will likely start school in Year 0 and transition to Year 1 the following February. This means that at the start of Year 2 of primary school, excluding holidays, children's potential time spent in school varies from approximately 4 to 11 months[3] (refer to Appendix Figure 3.A.1 for a graphical exposition). Another important characteristic of the school system is the school holiday period. There are four different school holiday periods in the New Zealand calendar year

[3] The maximum is 10.8 and the minimum 4.2, leading to a difference of 6.6 months. Details will follow.

(refer to Figure 3.1). The horizontal sections of the dark gray line (penroll) show the four holiday periods throughout the year and therefore show that there is not a linear relationship between age and the time spent in school. All the children born during these holidays start school at the same time on the first day of the new term. This gives us variation in age for students starting school at the same time (i.e. after the holidays). Therefore, the amount of time each child can potentially spend in school (maximum length of schooling) varies because of these characteristics and is not a linear function of his/her age. This is key for my identification strategy which follows previous work for the Netherlands by Leuven et al. (2010).

Figure 3.1 shows the relationship between a child's date of birth and his/her potential 'maximum length of time spent in school'. *Penroll* only includes teaching days while *Penroll0* also includes school holidays and weekends. The horizontal segments on *Penroll* reflect being born during school holidays and the segments with negative slope are for children born outside of school holidays. There are a total of four horizontal segments reflecting the four different periods of school holidays in a calendar year[4]. On my time axis, the first holidays are from July 9th to July 24th, the second from September 24th to October 9th, the third from December 20th to February 6th (which includes the Christmas and New Year holidays), and the fourth from April 14th to April 25th. Children who turn five on the same downward-sloping segment have a one to one relationship between the time potentially spent in school and age, i.e., an additional day of age leads to an additional day potentially spent in school. Any differences in the test scores of these children can be attributed to changes in their 'maximum length of schooling' as well as changes in their age (or randomly distributed changes in child/regional/parental characteristics). In comparison, children who turn five on the same horizontal segment (i.e., during a holiday period) all start school at the same time after the school holidays in the upcoming school term and so while they differ in age, they do not differ in the maximum time spent in school. Crucially, this allows us empirically to isolate the returns to time spent in school (mechanism C) from RAEs (mechanism B) – while absolute age effects (mechanism A) do not occur in a system where children start school at the same age.

[4] These holidays fall on slightly different days for some years used in my data. I have used the data for year 2005 for ease. The total number of holidays has not changed over the years.

3.4 NEW ZEALAND SCHOOL SYSTEM

Broadly speaking, there are three levels of the New Zealand education system:

1. From birth to school entry, known as early childhood education.
2. From Year 0 to Year 13 (about age 5-18), known as primary and secondary education.
3. Above Year 13 (from about age 18 onwards) – higher/tertiary and vocational education.

My study focuses on stage 2 above, i.e. the effects of early primary education on secondary school achievement and entry into tertiary education. Specifically, I examine the effects of differences in the initial time spent in primary school (due to differences in the dates of birth – as mentioned previously) on standardised achievement results near the end of high school.

New Zealand secondary schools operate a national qualification system known as the NCEA. This is what I use as one of the measures of standardised achievement, as described in detail below.

Another measure to assess the performance of a student, the second measure I use in this study, is known as UE. It is given to students based on specific NCEA results/achievements.

3.4.1 NATIONAL CERTIFICATE OF EDUCATIONAL ACHIEVEMENT (NCEA)

The NCEA are the primary national assessment tool for secondary school students in New Zealand (New Zealand Qualifications Authority, 2013/14). The New Zealand Qualifications Authority (NZQA) administers the NCEA and students do not have to apply to participate; they are automatically included.

NCEA qualifications are recognized by businesses, and used by colleges and universities both in New Zealand and abroad. Every student is assigned a unique identifier known as the National Student Number (NSN). The student or an employer/university can then use this unique number to search for the individual's NCEA results in an NZQA database.

NCEA tests the performance of students in various subjects, known as *standards*. For example, in mathematics standards, application of numeric thinking is measured. When students demonstrate a required level of knowledge/skills in a standard, they are awarded NCEA *credits*. Students need to obtain a specific number of credits in order to *achieve* an NCEA certification.

NCEA certification has three consecutive levels, based on the level of the evaluated knowledge/skills. Typically, students work through NCEA levels 1-3 in their secondary school Years 11-13, respectively. Receiving NCEA Merit or NCEA Excellence can officially recognize students' quality of work for a given level.

3.4.2 UNIVERSITY ENTRANCE (UE)

The minimum entrance requirement into a New Zealand university is UE. Gaining UE is the requirement of all New Zealand universities and some universities then have additional requirements beyond UE (Shui, 2017). The UE qualification is based on specific credits from NCEA levels 2 and 3 and is the minimum requirement for direct admission to a university in New Zealand (New Zealand Qualifications Authority, 2013/14).

To qualify for a UE, a student needs:

- An NCEA level 3 qualification;
- Approved subjects: 14 credits in each of three approved subjects[5] at NCEA level 3;
- A literacy requirement: 10 credits at NCEA level 2 or above, made up of 5 credits in reading and 5 credits in writing.
- A numeracy requirement: 10 credits at NCEA level 1 or above in relevant achievement standards; or all three numeracy standards (26623, 26626 and 26627).

Once a student has met the requirements for UE it will appear on his/her Record of Achievement.

3.4.3 SCHOOL DECILES

School decile is used in my study to control for socio-economic characteristics of the students and schools in my models. Later in Chapter 4, I use it as a proxy to identify the characteristics of the students and classify them in different groups to do robustness checks.

To explain the decile classification in brief, each school throughout New Zealand has been given a decile rating. It shows the socio-economic ranking of the census area sending children to each school. Decile 1 schools are the lowest ranked, implying that a high percentage of students in that particular school are from a low socio-economic background, while decile 10 schools are the highest ranked, implying that students in that particular school typically have a high socio-economic background. By design, each decile has approximately the same number

[5] A list of approved subjects is available here: https://www.nzqa.govt.nz/qualifications-standards/awards/university-entrance/approved-subjects/

of schools, i.e. roughly 10%. The decile rank is not in any way an indicator of the quality of education provided by the school.

Historically, the main objective of creating a decile ranking system was to determine how much disadvantage-related funding each state or state-integrated school should get. Schools in low deciles get the most funding per student. The New Zealand MOE re-calculates deciles every five years. The decile calculation is based on certain relative socio-economic factors of the community that students of a school come from. These factors include: household crowding; percentage of residents with income in the lowest twenty percent nationally; percentage of parents in low-skill occupational groups; percentage of parents without an educational qualification; and percentage of parents who are receiving income support benefits from the government.

For my analysis, I use the high school decile ranking for each student instead of the primary school decile ranking because of data limitations. Since the data used in the study is for children graduating or leaving school between 2009 and 2016 (details to follow), the primary school decile ranking for this population goes back to mid-late 1990s when adata availability was sparse.

3.5 DATA

The data used in this study is from the IDI provided by Stats NZ. The IDI is a large research database created by Stats NZ. It contains data about people and communities in the areas of education and training, income and work, benefits and social services, demographic information, tax, health, justice, housing etc. Data is compiled with the help of different government agencies and ministries, surveys conducted by Stats NZ, and some non-government organizations as well (refer to Figure 3.2).

The process of getting access to IDI is very well designed and organised. Stats NZ have set up secure data labs in different cities throughout New Zealand. Researchers who require access to the data need to go through a thorough application and training process. Specifically, a researcher has to first apply to get access to the data, providing a research proposal with a list of variables required. Stats NZ check this research proposal in detail, along with the applicant's CV and reports from two referees. Once a proposal is approved, the researcher has to go through a confidentiality-training programme. The whole process usually takes at least two months from data application to access.

The data used in this study contain information on each student who graduated or left a NZ secondary school between 2009 and 2016. For my analysis, I use the variables shown in Table 3.1.

Recall that for each student in high school (where we measure NCEA and UE achievement), we need to refer back to his/her fifth birthday and hence access to primary school education. As information about actual primary school enrolment date is very sparse for my older cohort (who turned five sometime between 1990 and 2000), this study uses *potential* enrolment (please refer back to Figure 3.1) instead of actual enrolment in school. More importantly, due to parents' choice in timing the start of school of their children (between turning 5 and 6), actual enrolment is likely to suffer from endogeneity. Potential enrolment would need to be used as an instrumental variable if actual enrolment were available. I use an intent-to-treat approach in its absence.

3.6 METHODS & RESULTS

Table 3.1 descriptively shows the characteristics of students in my sample. The MOE data in the IDI contains around 541,455 records on high school leavers. The first restriction I make is restricting the sample to those who left school because they had finished school (i.e. 'end of schooling') as I do not want to include students leaving school for other purposes such as to continue studies elsewhere in New Zealand or abroad. This restriction dropped around 80,000 records. The second restriction is to isolate only domestic students[6] as I only want students who started and finished school in New Zealand. Then, I check for duplicate observations in the data set and find 97 duplicate observations. I keep the latest data for those duplicates determined by comparing the students' recorded age, highest NCEA level, school leaving year, and the latest address. I also check for inconsistencies (e.g., a student with more than one gender recorded, students with abnormal dates of birth) and remove those individuals. After all the restrictions, I am left with 411,765 observations[7].

The population for my key analysis (Table 3.4, Table 3.5, and Table 3.6) is somewhat different than the publically available data provided by MOE on the education counts

[6] I identified domestic, New Zealand-born students using two different variables: i) One on the type of student - domestic, exchange and IFP and ii) the other on refugee status - New Zealand born, refugee, or migrant. I focus on 'domestic' and 'New Zealand born' students in analysis.

[7] All the numbers of observations reported here are very close to the exact values but not exactly the same. I do not report the exact numbers of observations because of the Stats NZ privacy clause.

website[8] for two reasons: 1. I restrict my analysis to domestic students only; 2. I include non-NCEA classification systems such as International Baccalaureate in my model as well by converting them to NCEA equivalent levels (refer to Appendix Table 3.A.1). For a detailed description of the difference in population, refer to Appendix Table 3.A.2.

I use the NCEA and UE achievement as outcome measures in my analysis. The exact date (rather than month) of birth would be ideal for the construction of the key explanatory variable, 'length of schooling' but unfortunately is not available in the IDI data set provided by Stats NZ[9]. In its absence, I randomly create the date of birth for each student based on information about his/her month (and year) of birth and I calculate the 'maximum length of time spent in school' accordingly. As mentioned previously, such measurement error in my key right-hand-side variable may bias any estimated effects downwards, making them conservative estimates of the returns to schooling. Later, in Chapter 4, I check the robustness of my results by re-doing my analysis selecting fixed alternate dates of birth - the 1st of each month, the 15th of each month, and the last of each month. The results (in Table 3.2) show that there is no substantial impact of this on my results.

3.6.1 EXOGENEITY OF THE MAXIMUM LENGTH OF SCHOOLING

Crucial to my analysis is the assumption that children's birth dates are not timed with the school calendar in mind and that parental characteristics do not systematically differ among children born at different points during the school year. In other words, I assume that the timing of the fifth birthday, and hence the maximum length of schooling, are exogenous. To test this, I estimate the following model:

$$penroll_{is} = \alpha + \gamma{*}X_{is} + s + t + r{*}t + \varepsilon_{is}.$$

Where *i* indexed a student in high school *s*. The *penroll* measures the amount of time spent in Years 0 and 1 of primary school; X is a vector of student characteristics including age and age^2 at the start of Year 2, gender, and ethnicity; *s* are high school fixed effects, *t* are year of birth dummies, and r*t are high school region*year of birth interactions.[10] The standard errors of all the regression models are corrected for clustering at the school level and are robust to heteroscedasticity. Table 3.3 shows the results.

---

8 https://www.educationcounts.govt.nz/statistics/schooling/senior-student-attainment/school-leavers2

9 Due to the privacy clause of data in the IDI, Stats NZ do not provide the exact date of birth of students to prevent revealing the identity of any individual. There are no exceptions to this rule.

10 Omitted Categories: Ethnicity=New Zealand European; Year of Birth=1988.

I expect to find no significance in any of the variables apart from the age variable to suggest that my results are exogenous. As expected, the exogeneity check shows the significance of age and age squared variable (refer to Figure 3.1). The check also shows significance in one of the ethnicity dummies. Surprisingly, the results suggest that being Māori decreases the *potential* amount of time spent in school compared to being New Zealand European. However, the effect is minute. For example, Māori children spend 0.003 less months – or 0.09 of a day less – in school than New Zealand Europeans.

The R-square of the above regression is 99%, which is no surprise as the age, and age square variables in the model are very closely, and mechanically, related to the maximum length of schooling (refer to Figure 3.1). When I remove the age and age squared variables from the model, the R-square drops down to approximately 2%. So as foreshadowed, it is reasonable to assume that New Zealand parents are not trying to select birthdate based on school start dates 5 years later.

3.6.2 INFLUENCE OF THE TIME SPENT IN SCHOOL ON LATER EDUCATIONAL OUTCOMES

I move next to my key analysis of the influence of the time spent in school on later educational outcomes measured by NCEA and UE results (Table 3.4 - Table 3.7). I run four different regressions for different levels of NCEA and UE:

1. NCEA1 = At least NCEA level 1 achieved (Table 3.4);
2. NCEA2 = At least NCEA level 2 achieved (Table 3.5);
3. NCEA3 = NCEA level 3 achieved (Table 3.6);
4. UE achieved (Table 3.7).

The models I run are given by:

$$NCEA1/NCEA2/NCEA3/UE_{is} = \alpha + \beta * penroll_{is} + \gamma * X_{is} + s + t + r*t + \varepsilon_{is}$$

I try two different approaches with my regressions: a linear probability model (LPM) and a probit model. Here, I only present the LPM with details to follow on why I choose this method over probit for my main analysis. The results of the probit model are available in the Appendix (Appendix Table 3.A.3, Table 3.A.4, Table 3.A.5, and Table 3.A.6). All regressions for these tables control for age and age squared, a gender dummy, six ethnicity dummies, thirteen year of birth dummies, year of birth and school region interaction dummies and school fixed effects. Standard errors are corrected for clustering at the school level and are robust to heteroscedasticity. An important thing to note here is that the population size for all four regressions is the same. For instance, even if some students drop out of school after achieving NCEA level 1, they are still part of my analysis for NCEA level 2, NCEA level 3, and UE and are considered as students who have not achieved these levels. Refer to the Appendix Figure 3.A.2.

The results in Table 3.4 show that an additional month of 'maximum time spent in school' results in an increase in achieving at least NCEA level 1 by 2.1 percentage points. This corresponds to about a 2.4% increase from the 89% baseline. Comparing the two extremes, being born in June rather than May increases the probability of achieving NCEA level 1 or above by 13.9 percentage points or 15.6%[11]. The ethnicity dummies show patterns similar to those found in the previous literature (Tofi, Flett, & Timutimu-Thorpe, 1996; Nakhid, 2003; Anae, Anderson, Benseman, & Coxon, 2002). On average, Māori and Pacific students are less likely to achieve NCEA level 1 than Asian or European students.

For NCEA level 2 (Table 3.5), an additional month of 'maximum time spent in school' results in a 3.7 percentage point increase in achievement from the 84% baseline. This makes it about a 4.4% increase. Comparing the two extremes as before, being born in June rather than May increases the probability of achieving NCEA level 2 or above by 24.4 percentage points or 29.1%.

There is an increasing impact of the 'maximum length of schooling' as children move up the NCEA levels. For NCEA level 3 (Table 3.6), an additional month of the 'maximum time spent in school' results in an increase in achieving NCEA level 3 by 3.9 percentage points, or 6.2% compared to a sample mean baseline at 63%. Again, comparing the two extremes, being born in June rather than May increases the probability of achieving NCEA level 3 by 25.7 percentage points or 40.9%. The impact for NCEA level 3 is the strongest among all NCEA levels.

The effects of early schooling on NCEA level 3 and UE are very similar. This is not too surprising given the importance of NCEA level 3 credits in being awarded UE. An additional month of the 'maximum time spent in school' results in an increase of 2.1 percentage points in the achievement of UE, compared to a sample mean baseline of 42% (Table 3.7). This equals to about a 4.9% increase compared to a 6.2% increase for NCEA level 3. When

[11] This is calculated by multiplying the effects of 1 month by 6.6; i.e., the difference between the maximum and the minimum potential time spent in school.

comparing the two extreme cases, being born in June rather than May (a 6.6 month difference) increases the probability of achieving UE by 13.9 percentage points or 32.6%.

There has been a debate (Hippel, 2015; Friedman, 2012; Hoetker, 2007; Caudill, 1988; Dubin & Rivers, 1989) about the relative merits of the linear regression model vs. a probit. I use the linear regression model over probit as the main model for my analysis following the approach from Friedman & Schady (2013) and since I have many fixed effects.

More pragmatically, my probit (or logit) model did not converge with the inclusion of school fixed effects and the large number of year*region interaction dummies.

I therefore run the following different models:

		A	B
		With year*region interaction dummies and school fixed effects	Without year*region interaction dummies and school fixed effects
1.	LPM	✓	✓
2.	Probit	X	✓

As explained above, I am unable to run model 2A. I therefore run 1B and 2B and compare the results of the two (refer to Appendix Table 3.A.3, Table 3.A.4, Table 3.A.5, and Table 3.A.6). The results show very similar marginal effects evaluated at mean values (and similar standards errors) for most of the variables. For instance, for UE, the coefficient on my main variable, penroll (maximum length of schooling), is 0.025 in the LPM compared to 0.030 in the probit, both highly statistically significant.

The reason for the parsimonious model being theoretically correct is the basic assumption of our analysis that the penroll variable, i.e. the time spent in school, is exogenous and hence orthogonal to student and school characteristics. Theoretically, even just raw correlations between high school achievement and penroll should suffice.

3.7 CONCLUSIONS

Due to the distinctive schooling system of New Zealand, in which children can begin school right after their fifth birthday, I am able to evaluate the effect of the potential length of schooling on NCEA and UE results, autonomous from the effect of age. Controlling for demographic and socio-economic characteristics, I find that increasing the maximum length of schooling substantially increases the probability of achieving NCEA and UE results. The magnitudes are shown in Table 3.8.

The 6.6 months category shows the two extreme cases, i.e., being born in early June rather than late May. Hence, the study strongly suggests that differences in the timing of birth – and hence in school start date – have large impacts on achievement of children even years later, in high school.

Chapter 3 again confirms, like Chapter 2, the effects of early schooling on later educational achievement. My replication in Chapter 2 and the original study by Leuven et al. (2010) suggests positive effects of early schooling for disadvantaged children with the replication also showing positive effects for non-disadvantaged children. This beneficial effect is observed only two years after the primary school start; i.e., around the age of six. Chapter 3 shows similar positive effects for students some 10-13 years after their school start; i.e., around the age of 15-18. Hence, early schooling seems to have large beneficial effects both in the short run and in the longer term.

Is it surprising to observe that additional schooling at an age of five or six can have so large effects on a child's educational achievement 10-13 years later? I believe it is not. Boucher and Mutimer (1994); Barnsley & Thompson (1988); Barnsley, Thompson, & Barnsley (1985); Daniel & Janssen (1987); Grondin, Deshaies, & Nault (1984); and Cobley et al. (2009) find long-term RAEs in professional sports. Clark et al. (2006) suggest that low expectation of success in children (that could have originated years ago) results in failure of a standard high school qualification, even after controlling for IQ and socio-economic characteristics. Lleras-Muney (2005) suggests that early school attendance may have long-term effects on mortality in the USA. My analysis fits in with this previous literature. The following chapter (Chapter 4) assesses the robustness – and the distribution – of these long term benefits.

Figure 3.1: The relationship between the maximum length of time spent in school and the date of birth for a given cohort

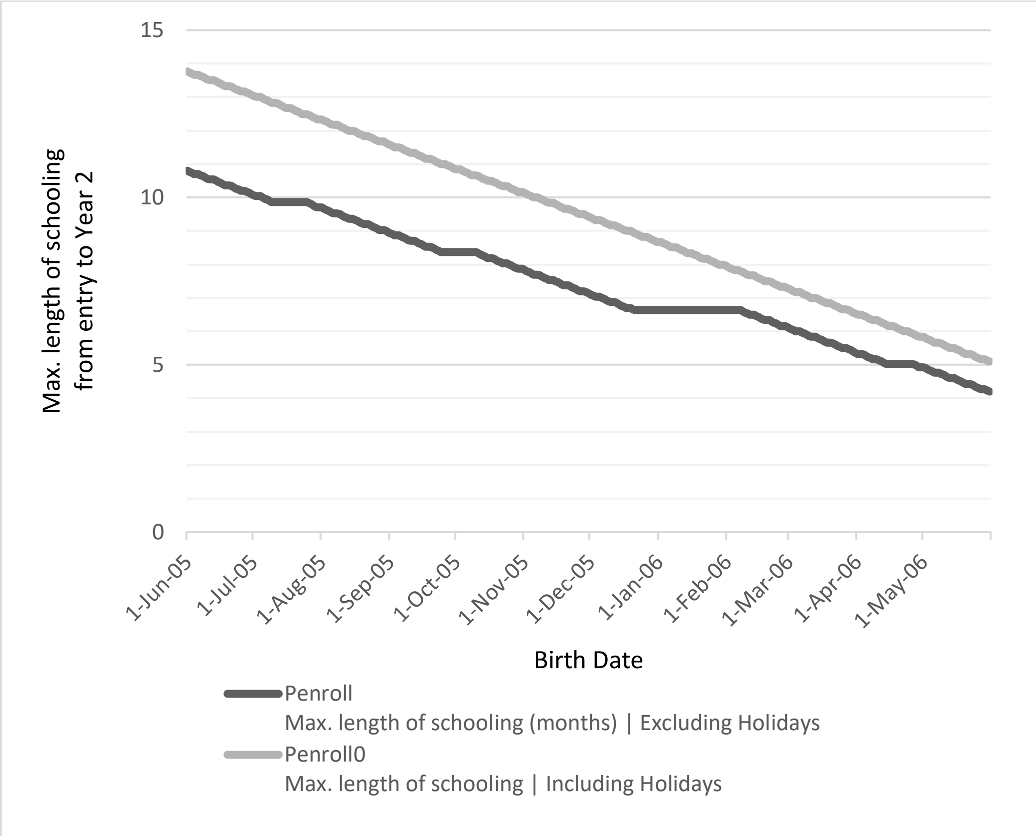
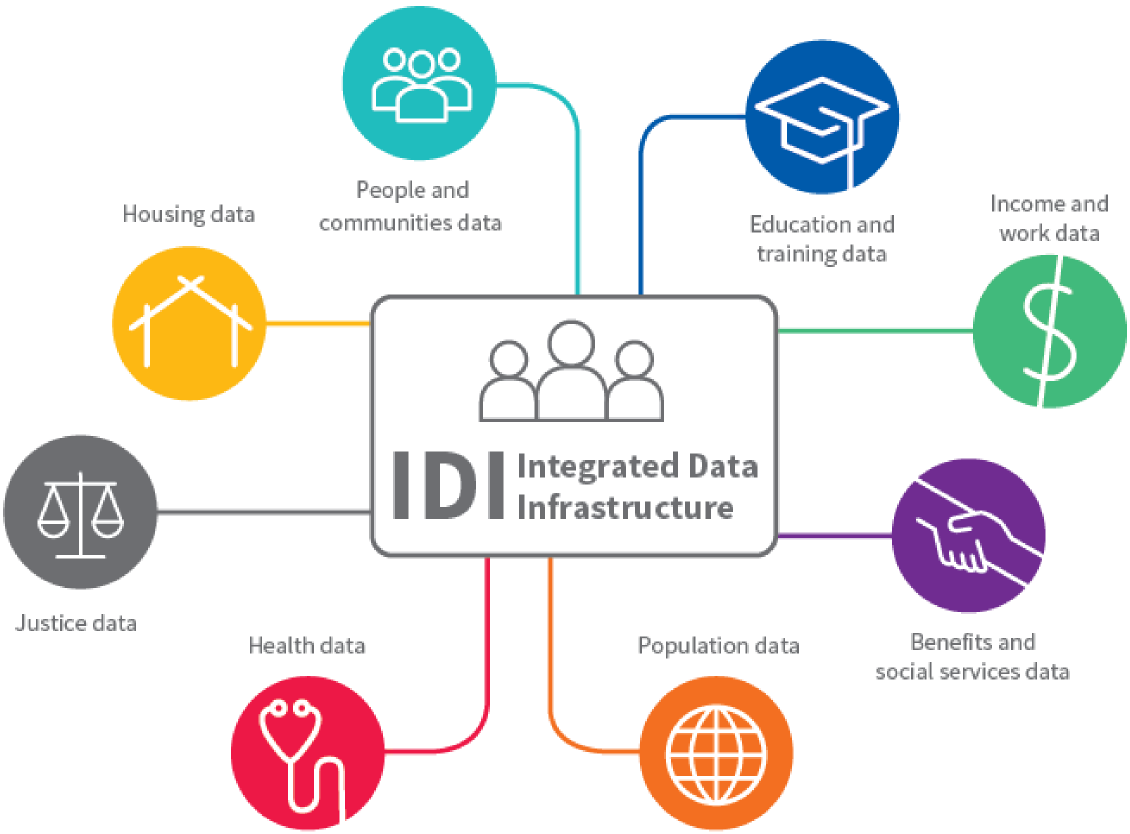


Figure 3.2: Data in the Integrated Data Infrastructure (IDI)



Source: (http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx), Integrated Data Infrastructure, Statistics New Zealand. Accessed on 28th November, 2018.

Table 3.1: Description of Variables

Name of Variable	Details
NCEA†	Categorical variable for the highest NCEA level achieved *(0 if NCEA not achieved; 1 if level 1 achieved; 2 if level 2 achieved; 3 if level 3 achieved; 99 for missing values)*
NCEA1+	At least NCEA level 1 achieved *(0/1)*
NCEA2+	At least NCEA level 2 achieved *(0/1)*
NCEA3	NCEA level 3 achieved *(0/1)*
UE	University Entrance achieved *(0/1)*
Penroll	Potential enrolment in months (time spent in school) without holidays – based on a random selection of birth date within a given month
Age m	Age in months at the start of Year 2 of school
Age m$_2$	Age in months - squared at the start of Year 2 of school
Female	Gender of the student *(0/1)*
Ethnicity	Ethnicity of the individual *(New Zealand European, Māori, Australian, European, Pacific People, Asian, Other ethnicity, Not stated)*
Dob y	Year of birth *(1988 to 2001)*
School decile	School deprivation decile *(1-10)*
School region	The region of the school (*Northland, Auckland, Waikato, Bay of Plenty, Gisborne, Hawkes Bay, Taranaki, Manawatu-Whanaganui, Wellington, West Coast, Canterbury, Otago, Southland, Tasman, Nelson, Marlborough*)

† This variable is not used in Chapter 3 but is used later in Chapter 4.

Table 3.2: Descriptive Statistics (mean values and standard deviations)

		Male	Female	NZ European	Māori	Asian	Australian	European	Pacific People	Total
NCEA1	*No. of observations*	163,458	167,862	206,310	58,737	21,867	2,028	14,916	21,471	331,320
	Mean	0.868	0.911	0.920	0.753	0.965	0.890	0.950	0.853	0.890
	Standard Deviation	0.339	0.284	0.271	0.432	0.185	0.312	0.218	0.355	0.313
NCEA2	*No. of observations*	163,458	167,862	206,310	58,737	21,867	2,028	14,916	21,471	331,320
	Mean	0.809	0.870	0.872	0.673	0.953	0.845	0.919	0.806	0.840
	Standard Deviation	0.393	0.336	0.334	0.469	0.212	0.362	0.273	0.396	0.367
NCEA3	*No. of observations*	163,458	167,862	206,310	58,737	21,867	2,028	14,916	21,471	331,320
	Mean	0.554	0.702	0.661	0.402	0.879	0.658	0.760	0.570	0.629
	Standard Deviation	0.497	0.457	0.473	0.490	0.327	0.475	0.427	0.495	0.483
UE	*No. of observations*	210,246	201,522	248,667	80,163	24,606	2,511	17,430	31,023	411,765
	Mean	0.356	0.496	0.478	0.185	0.743	0.473	0.589	0.250	0.425
	Standard Deviation	0.479	0.500	0.500	0.388	0.437	0.499	0.492	0.433	0.494
PENROLL	*No. of observations*	210,246	201,522	248,670	80,163	24,606	2,511	17,430	31,023	411,765
	Mean	7.461	7.465	7.468	7.461	7.438	7.432	7.449	7.448	7.463
	Standard Deviation	1.807	1.811	1.804	1.820	1.785	1.841	1.828	1.824	1.809
AGE M	*No. of observations*	210,246	201,522	248,670	80,163	24,606	2,511	17,430	31,023	411,765
	Mean	74.125	74.133	74.140	74.124	74.086	74.064	74.093	74.094	74.129
	Standard Deviation	S	S	S	S	S	S	S	S	S
FEMALE	*No. of observations*	-	-	248,670	80,163	24,606	2,511	17,430	31,023	411,765
	Mean	-	-	0.488	0.491	0.495	0.484	0.483	0.498	0.489
	Standard Deviation	-	-	0.500	0.500	0.500	0.500	0.500	0.500	0.500
DOB Y	*No. of observations*	210,246	201,522	248,670	80,163	24,606	2,511	17,430	31,023	411,765
	Mean	1995	1995	1995	1995	1995	1995	1995	1995	1995
	Standard Deviation	S	S	S	S	S	S	S	S	S
SCHOOL DECILE	*No. of observations*	200,721	189,072	234,972	74,925	23,805	2,379	16,578	30,426	389,796
	Mean	6.127	6.137	6.783	4.405	6.873	6.991	7.638	3.803	6.132
	Standard Deviation	2.611	2.660	2.249	2.460	2.584	2.381	2.081	3.483	2.635

Note: All figures have been randomly rounded to base 3 (RR3) – the number is randomly rounded to either the nearest base above or below the number – following the Stats NZ privacy requirement. Standard Deviations for AGE M and DOB Y have been supressed (S) due to a privacy clause. 'Age m' is Age in months at the start of Year 2.

Table 3.3: Exogeneity Check

Linear regression	Number of observations	411,765
	R-squared	0.9918
	Root MSE	0.16372
(Std. Err. adjusted for 561 clusters by school)		

Penroll	Coef.	Robust Std. Err.
Age m	-0.479***	0.004
Age m2	0.672***	0.003
Female	0.001	0.001
Ethnicity: Maori	-0.003***	0.001
Ethnicity: Asian	0.000	0.001
Ethnicity: Australian	-0.002	0.003
Ethnicity: European	0.003	0.001
Ethnicity: Pacific People	0.001	0.001

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 3.4: Effect of maximum schooling on achieving NCEA level 1

Linear regression	Number of observations	331,320
	R-squared	0.2135
	Root MSE	0.27794
(Std. Err. adjusted for 559 clusters by school)		

NCEA1	Coef.	Robust Std. Err.
Penroll	0.020***	0.003
Age m	0.038***	0.007
Age m2	-0.034***	0.005
Female	0.049***	0.003
Ethnicity: Maori	-0.127***	0.008
Ethnicity: Asian	0.017***	0.003
Ethnicity: Australian	-0.031***	0.007
Ethnicity: European	0.012**	0.004
Ethnicity: Pacific People	-0.050***	0.007

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 3.5: Effect of maximum schooling on achieving NCEA level 2

Linear regression	Number of observations	331,320
	R-squared	0.22
	Root MSE	0.3242
(Std. Err. adjusted for 559 clusters by school)		

NCEA2	Coef.	Robust Std. Err.
Penroll	0.037***	0.004
Age m	0.039***	0.007
Age m_2	-0.041***	0.005
Female	0.067***	0.003
Ethnicity: Maori	-0.147***	0.006
Ethnicity: Asian	0.040***	0.005
Ethnicity: Australian	-0.030***	0.008
Ethnicity: European	0.023***	0.004
Ethnicity: Pacific People	-0.050***	0.007

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 3.6: Effect of maximum schooling on achieving NCEA level 3

Linear regression	Number of observations	331,320
	R-squared	0.2401
	Root MSE	0.42161
(Std. Err. adjusted for 559 clusters by school)		

NCEA3	Coef.	Robust Std. Err.
Penroll	0.039***	0.004
Age m	0.020	0.010
Age m_2	-0.029***	0.007
Female	0.152***	0.008
Ethnicity: Maori	-0.177***	0.006
Ethnicity: Asian	0.139***	0.008
Ethnicity: Australian	-0.011	0.010
Ethnicity: European	0.059***	0.005
Ethnicity: Pacific People	-0.090***	0.009

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 3.7: Effect of maximum schooling on achieving UE

Linear regression	Number of observations	411,765
	R-squared	0.2595
	Root MSE	0.42576
(Std. Err. adjusted for 561 clusters by school)		

UE	Coef.	Robust Std. Err.
Penroll	0.021***	0.004
Age m	-0.003	0.009
Age m_2	-0.006	0.006
Female	0.138***	0.009
Ethnicity: Maori	-0.193***	0.009
Ethnicity: Asian	0.181***	0.009
Ethnicity: Australian	-0.013	0.010
Ethnicity: European	0.065***	0.006
Ethnicity: Pacific People	-0.187***	0.011

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 3.8: Effects of an increase in maximum schooling on NCEA and UE in terms of percentage points and percentages

	Sample Mean Baseline	Increase in maximum schooling by:	Effects in terms of percentage points (pp) and percentages (%)
NCEA level 1	89%	1 month	2.1 pp
		1 month	2.4%
		6.6 months	15.6%
NCEA level 2	84%	1 month	3.7 pp
		1 month	4.4%
		6.6 months	29.1%
NCEA level 3	63%	1 month	3.9 pp
		1 month	6.2%
		6.6 months	40.9%
UE	42%	1 month	2.1 pp
		1 month	4.9%
		6.6 months	32.6%

3.8 APPENDIX

Table 3.A.1: Highest attainment variable – classification table.

Code	Description	Percent	NCEA-Equivalent Measure
0	No Formal Attainment	2.18	NCEA not achieved
10	1 – 13 credits at level 1	1.11	NCEA not achieved
13	Other level 1 NQF Qualification	0.16	At least NCEA Level 1 achieved
14	NCEA level 1 not further defined	0.00	At least NCEA Level 1 achieved
15	NCEA level 1 achieved	3.47	At least NCEA Level 1 achieved
16	NCEA level 1 with merit	0.18	At least NCEA Level 1 achieved
17	NCEA level 1 with excellence	0.01	At least NCEA Level 1 achieved
20	1 – 13 credits at level 2	0.36	NCEA not achieved
24	NCEA level 2 not further defined	0.00	At least NCEA Level 2 achieved
25	NCEA level 2 achieved	15.77	At least NCEA Level 2 achieved
26	NCEA level 2 with merit	0.62	At least NCEA Level 2 achieved
27	NCEA level 2 with excellence	0.11	At least NCEA Level 2 achieved
30	1 – 13 credits at level 3	0.13	At least NCEA Level 1 achieved
33	Other level 3 NQF Qualification	0.49	NCEA Level 3 achieved
34	NCEA level 3 not further defined	0.00	NCEA Level 3 achieved
35	NCEA level 3 achieved	30.27	NCEA Level 3 achieved
36	NCEA level 3 with merit	12.23	NCEA Level 3 achieved
37	NCEA level 3 with excellence	4.65	NCEA Level 3 achieved
4	Other level 2 NQF Qualification	0.30	At least NCEA Level 2 achieved
40	3+ NZ Scholarships subjects	0.49	NCEA Level 3 achieved
43	National certificate at level 4	0.09	NCEA Level 3 achieved
51	14 – 39 credits at any level without level 1 literacy and numeracy credits	2.68	NCEA not achieved
52	14 – 39 credits at any level including level 1 literacy and numeracy credits	0.31	NCEA not achieved
53	40+ credits at any level without level 1 literacy and numeracy credits	2.21	NCEA not achieved
54	40+ credits at any level including level 1 literacy and numeracy credits	1.83	-----
55	30+ credits at level 2 or above	6.34	-----
56	30+ credits at level 3 or above	11.37	-----
60	International Baccalaureate Year 11	0.00	At least NCEA Level 1 achieved
61	International Baccalaureate Year 12	0.01	At least NCEA Level 2 achieved
62	International Baccalaureate Year 13	0.49	NCEA Level 3 achieved
70	Cambridge International Exams Year 11	0.04	At least NCEA Level 1 achieved
71	Cambridge International Exams Year 12	0.16	At least NCEA Level 2 achieved
72	Cambridge International Exams Year 13	1.87	NCEA Level 3 achieved

80	Accelerated Christian Education Year 11	0.02	At least NCEA Level 1 achieved
81	Accelerated Christian Education Year 12	0.01	At least NCEA Level 2 achieved
82	Accelerated Christian Education Year 13	0.02	NCEA Level 3 achieved
90	Other Overseas Awards Year 11	0.00	At least NCEA Level 1 achieved
91	Other Overseas Awards Year 12	0.00	At least NCEA Level 2 achieved
92	Other Overseas Awards Year 13	0.00	NCEA Level 3 achieved

Table 3.A.2: Difference in population of MOE and my analysis

	A	**B**	**C**
	MOE	My analysis: Missing NCEA excluded	My analysis: Missing NCEA included as NCEA not achieved
NCEA not achieved (0)	14%	11%	28%
At least NCEA 1 achieved (1)	86%	89%	72%

The table matches my data (columns B & C) with the MOE's publicly available data (column A) on the education counts website[12] and shows two different possibilities of using the missing values for the NCEA variable in my data. There are around 80,500 missing values for NCEA in my data set, which is approximately 19% of the total population. The table shows the difference between two scenarios: i) excluding these missing values from my analysis (column B) and ii) adding them to the category of NCEA not achieved (column C). It is evident that considering these values as missing and dropping them (column B) is more similar to the MOE's publicly available data (column A) so I choose this approach.

[12] https://www.educationcounts.govt.nz/statistics/schooling/senior-student-attainment/school-leavers2

Table 3.A.3: NCEA level 1 – OLS vs PROBIT

NCEA level 1						
Linear regression	Number of observations 331,320 R-squared 0.0776 Root MSE 0.30066 (Std. Err. adjusted for 559 clusters in schoolno)			**Marginal effects after probit** y = Pr(ncea1) (predict) = 0.90725632		
NCEA1	**Coef.**	**Robust Std. Err.**		**NCEA1**	**dy/dx**	**Std. Err.**
Penroll	0.017***	0.004		**Penroll**	0.015***	0.003
Age m	0.042***	0.009		**Age m**	0.038***	0.008
Age m2	-0.036***	0.006		**Age m2**	-0.032***	0.006
Female	0.041***	0.005		**Female**	0.040***	0.005
Ethnicity: Māori	-0.162***	0.013		**Ethnicity: Māori**	-0.160***	0.013
Ethnicity: Asian	0.041***	0.011		**Ethnicity: Asian**	0.052***	0.013
Ethnicity: Australian	-0.030***	0.008		**Ethnicity: Australian**	-0.035***	0.009
Ethnicity: European	0.028***	0.006		**Ethnicity: European**	0.034***	0.006
Ethnicity: Pacific People	-0.067***	0.011		**Ethnicity: Pacific People**	-0.075***	0.013

Table 3.A.4: NCEA level 2 – OLS vs PROBIT

NCEA level 2						
Linear regression	Number of observations 331,320 R-squared 0.0893 Root MSE 0.34992 (Std. Err. adjusted for 559 clusters in schoolno)			**Marginal effects after probit** y = Pr(ncea2) (predict) = 0.85872728		
NCEA2	**Coef.**	**Robust Std. Err.**		**NCEA2**	**dy/dx**	**Std. Err.**
Penroll	0.033***	0.004		**Penroll**	0.032***	0.004
Age m	0.045***	0.008		**Age m**	0.043***	0.008
Age m2	-0.043***	0.006		**Age m2**	-0.041***	0.006
Female	0.058***	0.008		**Female**	0.059***	0.008
Ethnicity: Māori	-0.191***	0.012		**Ethnicity: Māori**	-0.190***	0.012
Ethnicity: Asian	0.076***	0.013		**Ethnicity: Asian**	0.091***	0.015
Ethnicity: Australian	-0.028**	0.009		**Ethnicity: Australian**	-0.031**	0.010
Ethnicity: European	0.044***	0.007		**Ethnicity: European**	0.051***	0.007
Ethnicity: Pacific People	-0.066***	0.014		**Ethnicity: Pacific People**	-0.072***	0.015

Table 3.A.5: NCEA level 3 – OLS vs PROBIT

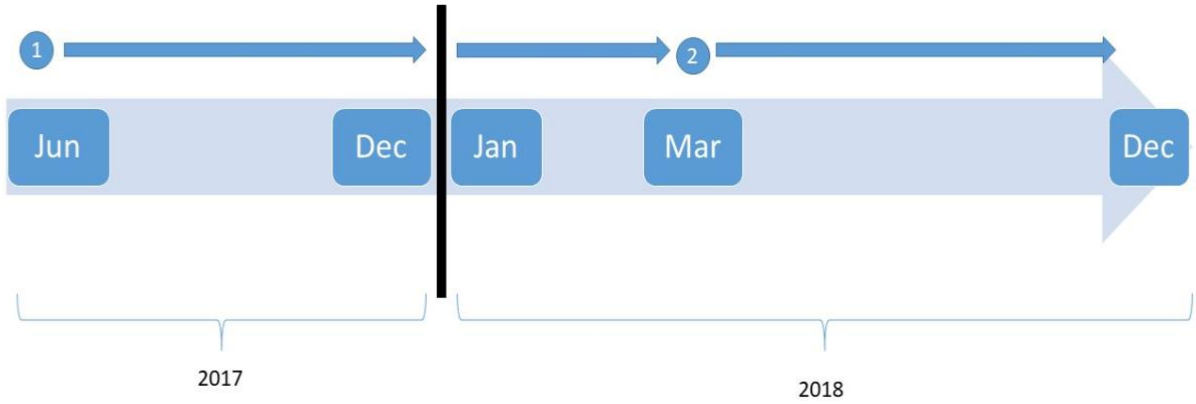
NCEA level 3					
Linear regression (Std. Err. adjusted for 559 clusters in schoolno)	Number of observations 331,320 R-squared 0.1132 Root MSE 0.45496		**Marginal effects after probit** y = Pr(ncea3) (predict) = 0.63681061		
NCEA3	**Coef.**	**Robust Std. Err.**	**NCEA3**	**dy/dx**	**Std. Err.**
Penroll	0.037***	0.006	**Penroll**	0.040***	0.006
Age m	0.028*	0.011	**Age m**	0.031**	0.012
Age m2	-0.033***	0.008	**Age m2**	-0.037***	0.009
Female	0.143***	0.014	**Female**	0.154***	0.015
Ethnicity: Māori	-0.249***	0.010	**Ethnicity: Māori**	-0.259***	0.010
Ethnicity: Asian	0.211***	0.015	**Ethnicity: Asian**	0.237***	0.016
Ethnicity: Australian	-0.002	0.012	**Ethnicity: Australian**	-0.003	0.013
Ethnicity: European	0.096***	0.008	**Ethnicity: European**	0.104***	0.008
Ethnicity: Pacific People	-0.091***	0.018	**Ethnicity: Pacific People**	-0.097***	0.020

Table 3.A.6: UE – OLS vs PROBIT

UE					
Linear regression (Std. Err. adjusted for 561 clusters in schoolno)	Number of observations 411,765 R-squared 0.1246 Root MSE 0.46249		**Marginal effects after probit** y = Pr(ue) (predict) = 0.40923673		
UE	**Coef.**	**Robust Std. Err.**	**UE**	**dy/dx**	**Std. Err.**
Penroll	0.019**	0.006	**Penroll**	0.022**	0.007
Age m	0.003	0.010	**Age m**	0.004	0.011
Age m2	-0.009	0.007	**Age m2**	-0.011	0.008
Female	0.138***	0.017	**Female**	0.150***	0.018
Ethnicity: Māori	-0.287***	0.011	**Ethnicity: Māori**	-0.294***	0.011
Ethnicity: Asian	0.261***	0.017	**Ethnicity: Asian**	0.278***	0.017
Ethnicity: Australian	-0.002	0.011	**Ethnicity: Australian**	-0.001	---
Ethnicity: European	0.110***	0.009	**Ethnicity: European**	0.113***	0.009
Ethnicity: Pacific People	-0.229***	0.019	**Ethnicity: Pacific People**	-0.222***	0.019

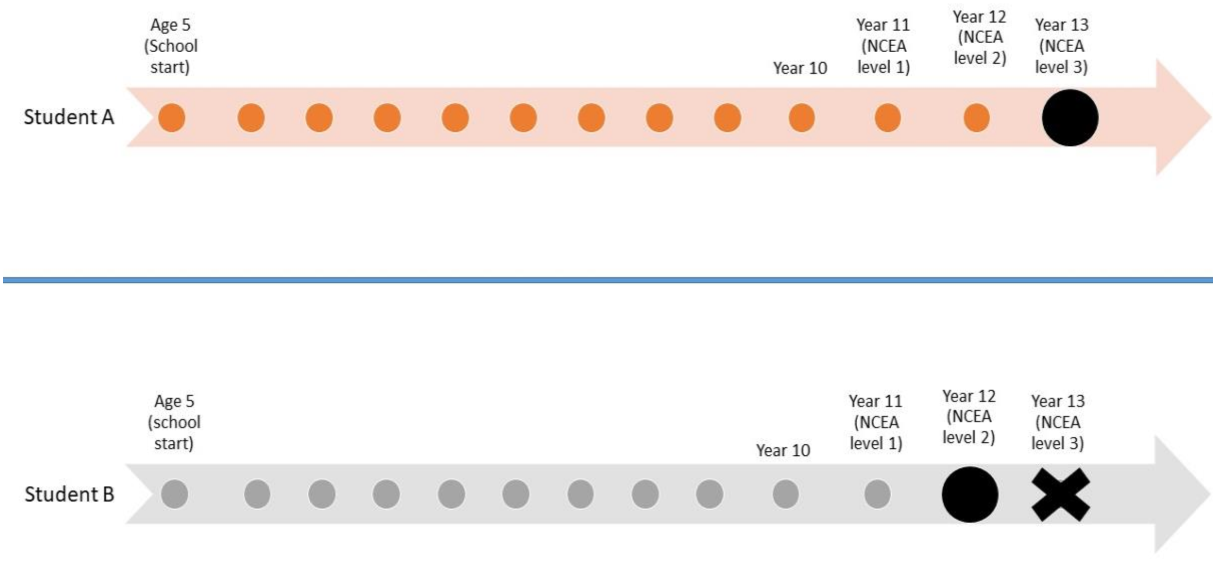
Note: All regressions in Tables A2-A6 also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. The marginal effects for the probit models are calculated at mean values.

Figure 3.A.1: Comparison of two hypothetical students starting school at different times



The diagram shows the example of two students labelled with 1 and 2 in calendar years 2017 & 2018 and how much time they will spend in school before the start of Year 2. Take student 1, who starts school in June 2017. This student will spend the rest of year 2017 (June – December) and the entire next year, 2018 (January – December) in school before he/she starts Year 2. Now take student 2, who starts school in March 2018. This student will only spend March till December of 2018 in school and will move on to Year 2 next year. Comparing the two students, student 1 has spent roughly 18 months in school whilst student 2 has spent only 10 months in school before starting Year 2. Taking this even further, another student born in May of 2018 will only spend 6 months in school before Year 2.

Figure 3.A.2: Comparison of two hypothetical students achieving different NCEA levels

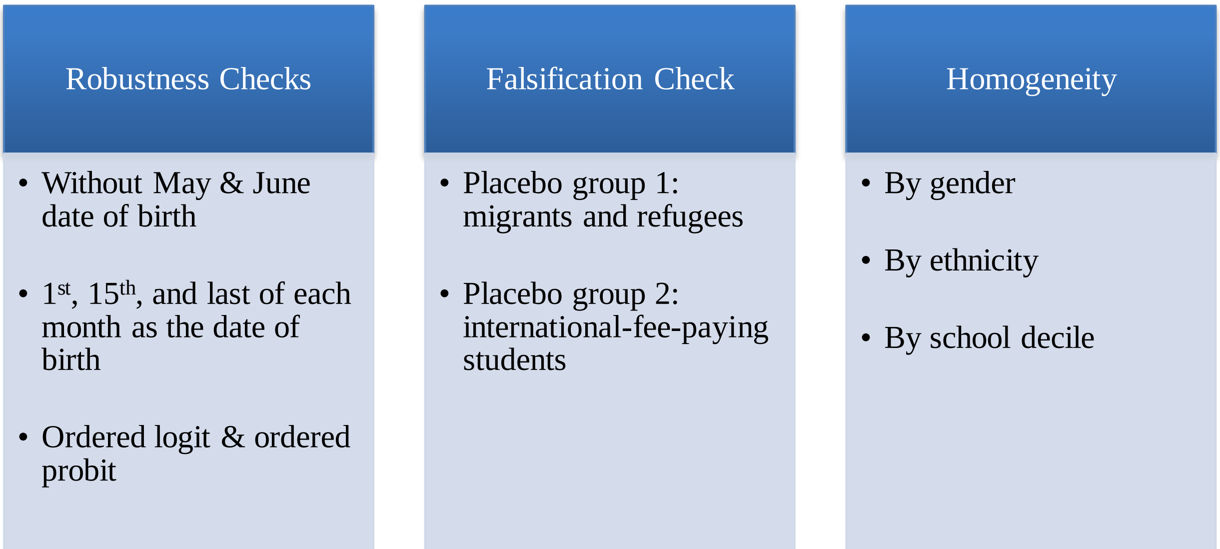


The diagram above gives an example of two different students (Student A and Student B) each achieving different NCEA levels. The small dots show the time spent in school. The large black dot shows the last year of school or last NCEA achievement level. The black cross shows the year or NCEA level not attended/not achieved. If I compare the two students, I can see that student A attended all 13 years and achieved NCEA level 3 whereas student B achieved only NCEA level 2 (either because he/she did not attend Year 13 or did but did not achieve NCEA level 3). When analysed in a model of NCEA level 3 achievement, both of these students are in the sample, with student A achieving level 3 and student B not achieving/attending level 3. The important thing to note here is that the sample size for different models – NCEA level 1, level 2, and level 3 (and UE) is the same. In other words, drop-outs are treated as non-achievers along with those who attempted the assessment but did not succeed in it.

CHAPTER 4. ROBUSTNESS, FALSIFICATION AND HOMOGENEITY CHECKS

4.1 INTRODUCTION

This chapter extends the results of Chapter 3 in several ways. In particular, it consists of robustness checks, falsification checks, and sub-samples analyses of the main results from Chapter 3. The following diagram shows the different checks I conduct in this chapter.



The descriptive statistics about the population of these groups are given in Table 4.1 and Table 4.2. The results of these checks generally endorse the findings of the main analysis conducted in Chapter 3.

4.2 ROBUSTNESS CHECKS

Three different types of checks have been conducted to evaluate the robustness of the main analysis:

1. Excluding the ‘extreme’ dates of birth in May & June;
2. Using different assumptions about the exact date of birth; and
3. Analysing educational achievement with ordered logit & probit.

4.2.1 DATES OF BIRTH WITHOUT MAY & JUNE

A key part of my main analysis is the date of birth of each individual. Since I use potential time spent in school as my key control variable with the assumption that everyone starts school on his/her 5th birthday, the date of birth has been used as a proxy for the time spent in school before the start of Year 2 of primary school. This makes the exact date of birth crucial for the main analysis.

As mentioned in Chapter 3, the extreme cases with the largest difference in *penroll* – time spent in school – are students with birth dates in May vs. June (reference Figure 3.A.1). A child born in late May will potentially spend approximately 6.6 months less in school than one born in early June. However, in these extreme cases, the correspondence between *potential* schooling and *actual* schooling is likely to be the weakest. Specifically, the parents or the school of a child born in May can decide to postpone actual school start and then place the child into Year 0 rather than directly into Year 1. Similarly, a child born in early June may be put into Year 1 in some schools. To account for these discrepancies, I remove all the observations that have a May or June birth date from my analysis as a robustness check.

After removing these observations from the data set, I conduct the same exogeneity test as done for the main analysis (Table 4.3). As expected, I find significance of the age and age squared variable. Similar to the results of the exogeneity test of my main analysis (Table 3.3) the results here also show significance of the Māori ethnicity dummy. This means that being a Māori decreases the *potential* amount of time spent in school compared to being a New Zealand European. However, the effect is minute as it was in the main analysis. It shows that Māori children spend 0.004 months – or 0.10 of a day – less in school than New Zealand European students.

The second step is the estimation of the same model as the main analysis for different NCEA levels as well as UE (Table 4.4). Both the results of the main analysis and the current analysis without students born in May and June are shown in the tables for an ease of comparison. The left hand side of the tables shows the results for the main analysis and the right hand side shows the robustness check.

The results of the robustness check show that an additional month of the 'maximum time spent in school' results in an increase in achieving NCEA level 1 by 1.2 percentage points, NCEA level 2 by 2.8 percentage points, NCEA level 3 by 3.4 percentage points and UE by 2.2 percentage points. Hence, removing students born in May or June from the analysis does not qualitatively change my main findings.

4.2.2 ALTERNATIVE PROXIES FOR THE DATE OF BIRTH

As mentioned above, the date of birth is an important variable in my analysis. The exact date of birth is unfortunately unavailable in the IDI data set (only month and year of birth are) and so dates of birth have been randomly assigned within each month (as explained in Chapter 3). To test the sensitivity of my results to this 'noise', all the models are re-estimated using alternative assumptions about the exact date of birth. Specifically, three alternative dates of birth are assigned to each individual:

1. The 1st of the month of birth;
2. The 15th of the month of birth;; and
3. The last (28th/30th/31st) of the month of birth.

As the imputed potential length of schooling decreases (e.g., the 15th vs. the 1st), holding later outcomes constant, I would expect to estimate higher returns per month. However, qualitatively, the main results should be fairly similar.

I again start with the exogeneity check (Table 4.5) of the maximum length of schooling followed by the effect of time spent in school on NCEA level 1, NCEA level 2, NCEA level 3, and UE separately.

Table 4.6 shows the results compared with the main analysis. When using a random date of birth, an additional month of schooling increases NCEA level 1 achievement by 2.1 percentage points. In contrast, assuming all students are born on the 1st of each month shows an increase of NCEA level 1 by 1.4 percentage points. Using the 15th of each month results in an increase of 2.3 percentage points while using the last of each month an increase of 2.9 percentage points. As expected, the estimated monthly return increases as the assumed time spent in school decreases. Similarly for NCEA level 2 achievement, the main model predicted an increase of 3.7 percentage points per month of potential early schooling while models with the 1st, 15th and last of each month show increases of 2.7, 3.9 and 4.4 percentage points, respectively. The effect of an additional month of schooling on NCEA level 3 was 3.9 percentage points in the main analysis and becomes 2.5, 4.0, and 4.5 percentage points for the 1st, 15th and the last of each month, respectively. Overall, while the estimated returns to early schooling depend on assumptions about the exact date of birth as expected, all of the estimated effects are significantly greater than zero and substantial.

4.2.3 ORDERED LOGIT/PROBIT

As a third robustness check, I estimate ordered logit and ordered probit models and compare the results of these with my main analysis. Although a direct comparison is difficult, I expect to find similar patterns when using binary and categorical achievement measures.

Recall Table 3.A.3, Table 3.A.4, Table 3.A.5, and Table 3.A.6 in Chapter 3, where I compare the results of the probit (or logit) model with the LPM. One of the reasons I choose the LPM over probit/logit is that the latter does not converge with the inclusion of school fixed effects and the large number of year*region interaction dummies. The ordered logit and the ordered probit models face the same difficulties, so I estimate a parsimonious model as follows:

$$NCEA\ achievment_{is} = \alpha + \beta * penroll_{is} + \gamma * X_{is} + s + t + r*t + \varepsilon_{is}$$

Here NCEA is now a categorical variable ranging from 0 to 3 and indicating the *highest* NCEA level achieved: none, NCEA level 1, level 2, or level 3.

Table 4.7 shows the marginal effects (evaluated at the mean values) of the ordered logit and ordered probit. The results of the two are fairly similar, as expected. The results of the ordered logit show that with one additional month of schooling, an individual is 1.7 percentage points less likely to be not achieving any NCEA level (i.e. NCEA=0), 0.7 percentage points less likely to be achieving at most NCEA level 1, 2.3 percentage points less likely to be achieving NCEA level 2 (but not level 3), and 4.7 percentage points more likely to be achieving NCEA level 3. The results of the ordered probit model (Table 4.7) show similar results.

It is hard to quantitatively compare these results with the LPM but they all point towards a similar pattern. Specifically, the LPM shows that the effects are the largest as children move towards the highest NCEA levels and the ordered logit/probit reinforce this finding.

4.3 FALSIFICATION CHECKS

The population of the main analysis, in Chapter 3, is restricted to domestic and New Zealand born students as explained previously (see footnote 6). This restriction is intended to exclude students who had started primary schooling outside of New Zealand (i.e., under a different primary school start policy) and who later came to New Zealand and took the NCEA tests. This restriction is important as our analysis relies on students having come through the rolling admission system and hence experiencing differences in the time spent in primary school before the start of Year 2.

As a falsification check, the same analysis is now conducted on children who are expected to have started schooling outside of New Zealand (i.e., have not gone through the same rolling admission system) but have later taken the NCEA tests. In this analysis, I expect to find no relationship between the maximum length of schooling and achievement on the NCEA and UE tests.

Two falsification checks have been conducted on two different student populations: i) migrants and refugees and ii) international-fee-paying students (Table 4.8 and Table 4.9 identify these placebo populations as well as the main population of interest in Chapter 3).

4.3.1 PLACEBO 1: MIGRANTS AND REFUGEES

The IDI dataset reports the migration status of each individual, distinguishing between: migrants, refugees, and New Zealand born students. In my first placebo test, I focus on migrants and refugees (whether they are considered domestic or international fee paying, IFP, as shown in Table 4.8). The assumption is that these students - even though most of them are now considered domestic – will not have started schooling in New Zealand.

Table 4.10 shows the ethnic composition of the population used for this placebo test. Approximately 15% are Indian, 15% Chinese, 13% Samoan, and 8% are Japanese. As Table 4.11 shows, migrants and refugees are divided evenly among the different school deciles. A high percentage of these students, 56% compared to 51% of New Zealand born students, achieve NCEA level 3 and about 46% achieve UE (Table 4.12 and Table 4.13).

Table 4.14 shows the exogeneity check. As expected, the migrant/refugee community does not benefit from the New Zealand primary school start policy in the way that domestic students do (Table 4.15). All of the estimated coefficients for NCEA and UE achievement are close to zero and many have a negative sign. This endorses the credibility of my main analysis.

4.3.2 PLACEBO 2: INTERNATIONAL-FEE-PAYING STUDENTS

The population for my second placebo test consists of only international-fee-paying students (refer to Table 4.9). These students are almost always (in 98% of cases) not classified as migrants, refugees or NZ born.

The population in this test consists of students from many different ethnicities but a large portion is from North East Asian countries (refer to Table 4.10) with around 47% being Chinese, 15% Korean, and 8% Japanese. Many of these students attend high decile schools with approximately 35% in decile 7 schools, 10% in decile 8 schools, 18% in decile 9 schools, and 13% in decile 10 schools (Table 4.11). These students have low NCEA achievement rates compared to New Zealand born students. Specifically, around 42% achieve NCEA level 3 and 38% achieve UE, compared to 51% and 42% among New Zealand born students, respectively (Table 4.12 and Table 4.13).

Unlike migrants and refugees, international-fee-paying students do not appear very different from the main study population in the placebo test (Table 4.15). Specifically, the potential time spent in school (had they started primary school in New Zealand) is a positive, large, and sometimes statistically-significant predictor of their high school achievement results[13]. I hope to explore this puzzling finding thoroughly in future work.

4.4 HETEROGENEITY OF EFFECTS

My next aim is to explore whether the beneficial effects of early schooling occur broadly or whether they are more concentrated in certain socio-demographic groups. Specifically, the potential heterogeneity of effects is investigated by three different dimensions:

1. Gender;
2. Ethnicity; and
3. School decile.

4.4.1 BY GENDER

The first homogeneity test conducted is by gender. I check whether there are any differences in the returns to early schooling between girls and boys. Male and female students have similar results for the exogeneity check, suggesting that the potential length of schooling is randomly distributed for both (Table 4.16). The causal effects of early schooling (Table 4.17) are somewhat larger for male students than for female students. Compared to the main model in which an additional month of the maximum time spent in school resulted in an increase in achieving at least NCEA level 1 by 2.1 percentage points and at least NCEA level 2 by 3.7 percentage points, the results here show increases of 1.8 and 2.7 percentage points for female students and 2.5 and 4.7 percentage points for male students. The results for NCEA level 3 and UE were 3.9 and 2.1 percentage points for the main analysis compared to 3.7 and 2.0 percentage points for females and 4.1 and 2.2 percentage points for males.

To summarise, I observe large and positive effects of early schooling for both genders, but especially among male students who experience larger benefits in absolute terms as well as relative to their (lower) mean performance.

4.4.2 BY ETHNICITY

The second homogeneity test is by ethnicity, focusing on the three largest groups in New Zealand:

1. New Zealand European;
2. Māori; and

[13] The population has 25% of observations from one particular school. Dropping this school from the analysis does not qualitatively change the results.

3. Asian.

Table 4.18 shows the exogeneity check and Table 4.19 shows the comparison of results between the different ethnic groups and the main model. New Zealand European students experience small effects of early schooling on NCEA levels 1 and 2 and the effects increase for NCEA level 3 and UE. On the other hand, students from a Māori background experience a relatively large effect on NCEA levels 1 and 2 but smaller effects as they move further to NCEA level 3 and UE. For Asian students, the effect is minimal throughout.

Overall, early school attendance seems to have the largest benefits for Māori students, followed by New Zealand Europeans, and – only weakly – Asians.

4.4.3 BY SCHOOL DECILE

The final test for homogeneity is conducted by high school decile, categorized into three different groups as follows:

1. Decile group 1 – high school deciles 1-4;
2. Decile group 2 – high school deciles 5-7; and
3. Decile group 3 – high school deciles 8-10.

Table 4.20 shows the exogeneity check. It is interesting to note that the exogeneity check holds even for parents of kids in high decile schools, since if anyone might try to time births for educational advantage it would likely be these (most educated) parents. Table 4.21 shows that students who study in low decile schools tend to do better in the lower NCEA levels (1 and 2) with increased early schooling. For instance, students from deciles 1-4 experience an increase of 2.8 percentage points at NCEA level 1 compared to 2.1 percentage points for the main analysis. At higher NCEA levels, their benefits become smaller. Specifically, the effect at NCEA level 2 is fairly similar to the main analysis (3.7 vs. 3.4 percentage points) and the effect at NCEA level 3 is much smaller (1.6 vs. 3.9 percentages points). On the other hand, the effect of early schooling on students in decile 8-10 schools is much smaller at lower NCEA levels (e.g., 1.1 percentage points vs. 2.1 percentage points at NCEA level 1) but it increases as they move to a higher NCEA level (3.6 percentage points vs. 3.9 percentage points at NCEA level 3).

The largest benefits of early schooling occur among students in high school deciles 5-7, rather than low-decile or top-decile schools. One interpretation of these findings is that low decile schools provide valuable early formal education but are constrained by own resources and/or (the lack of) parental effort to complement/endorse school activities at home (Ali and Menclova, 2018). At the other end of the spectrum, children from high decile schools may be using the school environment and in-home learning as substitutes (Leuven et al., 2010).

4.5 CONCLUSION

This chapter investigates the robustness of the analysis in Chapter 3 by conducting several robustness, falsification and heterogeneity checks. First, I check the robustness of my analysis by removing students with May and June dates of birth, using different dates of birth as proxy, trying ordered logit and ordered probit models. All these different checks show that my results are robust. Second, I run a falsification check with two different placebo populations. The results tentatively suggest that my results are robust. Finally, I explore whether the beneficial effects of early schooling occur broadly or whether they are more concentrated in certain socio-demographic groups. The results show that the effects are the strongest among male, Māori, and decile 5-7 students, followed by decile 1-4 students.

Table 4.1: Descriptive Statistics (mean values and standard deviations)

		Decile Group 1	Decile Group 2	Decile Group 3	1st of each month	15th of each month	Last of each month	Placebo Group 1	Placebo Group 2	w/o May & June
NCEA1	*No. of observations*	85,314	110,070	118,962	331,320	331,320	331,320	27,225	11,997	276,933
	Mean	0.826	0.905	0.966	0.890	0.890	0.890	0.904	0.596	0.891
	Standard Deviation	0.379	0.293	0.181	0.313	0.313	0.313	0.295	0.491	0.311
NCEA2	*No. of observations*	85,314	110,073	118,962	331,320	331,320	331,320	27,228	12,000	276,933
	Mean	0.761	0.850	0.940	0.840	0.840	0.840	0.881	0.582	0.841
	Standard Deviation	0.427	0.357	0.238	0.367	0.367	0.367	0.324	0.493	0.366
NCEA3	*No. of observations*	85,314	110,070	118,962	331,320	331,320	331,320	27,225	12,000	276,933
	Mean	0.500	0.603	0.795	0.629	0.629	0.629	0.717	0.495	0.630
	Standard Deviation	0.500	0.489	0.404	0.483	0.483	0.483	0.450	0.500	0.483
UE	*No. of observations*	115,038	138,183	136,575	411,765	411,765	411,765	35,010	14,157	344,214
	Mean	0.257	0.394	0.639	0.425	0.425	0.425	0.460	0.384	0.425
	Standard Deviation	0.437	0.489	0.480	0.494	0.494	0.494	0.498	0.486	0.494
PENROLL	*No. of observations*	115,038	138,183	136,575	411,765	411,765	411,765	35,010	14,157	344,214
	Mean	7.452	7.469	7.474	7.738	7.471	7.203	7.439	7.438	7.458
	Standard Deviation	1.817	1.809	1.802	1.809	1.806	1.792	1.815	1.812	1.483
AGE M	*No. of observations*	115,038	138,183	136,575	411,765	411,765	411,765	35,010	14,157	344,214
	Mean	74.107	74.141	74.151	74.620	74.153	73.639	74.075	74.073	74.145
	Standard Deviation	S	S	S	S	S	S	S	S	S
FEMALE	*No. of observations*	115,038	138,180	136,575	411,765	411,765	411,765	35,010	14,157	344,214
	Mean	0.484	0.481	0.490	0.489	0.489	0.489	0.475	0.479	0.489
	Standard Deviation	0.500	0.500	0.500	0.500	0.500	0.500	0.499	0.500	0.500
DOB Y	*No. of observations*	115,038	138,180	136,575	411,765	411,765	411,765	35,010	14,157	344,214
	Mean	1995.012	1995.052	1995.010	1994.995	1994.995	1994.995	1994.778	1994.287	1994.995
	Standard Deviation	S	S	S	S	S	S	S	S	S
SCHOOL DECILE	*No. of observations*	115,038	138,180	136,575	411,765	411,765	411,765	35,010	14,157	344,214
	Mean	2.777	6.114	8.975	11.087	11.087	11.087	7.226	9.042	11.083
	Standard Deviation	1.110	0.752	0.796	21.028	21.028	21.028	12.933	12.713	21.006

Table 4.2: Descriptive Statistics (Number of observations and percentages)

		Decile Group 1	Decile Group 2	Decile Group 3	1st of each month	15th of each month	Last of each month	Placebo Group 1	Placebo Group 2	w/o May & June
NEW ZEALAND EUROPEAN	*No. of observations*	42,621	95,697	96,648	248,664	248,664	248,664	726	---	208,281
	Percentage	37.05%	69.26%	70.76%	60.39%	60.39%	60.39%	2.07%	---	60.51%
MĀORI	*No. of observations*	42,654	22,077	10,197	80,166	80,166	80,166	63	---	66,765
	Percentage	37.08%	15.98%	7.47%	19.47%	19.47%	19.47%	0.18%	---	19.40%
ASIAN	*No. of observations*	5,811	6,186	11,808	24,606	24,606	24,606	19,521	11,949	20,766
	Percentage	5.05%	4.48%	8.65%	5.98%	5.98%	5.98%	55.76%	84.40%	6.03%
AUSTRALIAN	*No. of observations*	420	864	1,095	2,511	2,511	2,511	18	---	2,079
	Percentage	0.37%	0.63%	0.80%	0.61%	0.61%	0.61%	0.05%	---	0.60%
EUROPEAN	*No. of observations*	1,641	5,334	9,606	17,430	17,430	17,430	1,791	993	14,478
	Percentage	1.43%	3.86%	7.03%	4.23%	4.23%	4.23%	5.12%	7.01%	4.21%
PACIFIC PEOPLE	*No. of observations*	20,556	5,493	4,290	31,023	31,023	31,023	8,523	522	25,740
	Percentage	17.87%	3.98%	3.14%	7.53%	7.53%	7.53%	24.34%	3.69%	7.48%

Note: For both Table 4.1 and Table 4.2, decile group 1 consists of school deciles 1 to 4, decile group 2 consists of school deciles 5 to 7, and decile group 3 consists of school deciles 8 to 10. Placebo group 1 consists of migrant and refugee students and group 2 consists of IFP students. All figures have been randomly rounded to base 3 (RR3) – the number is randomly rounded to either the nearest base above or below the number – following the Stats NZ privacy requirement. In Table 4.1, 'Age m' is Age in months at the start of Year 2. Standard deviations for AGE M and DOB Y have been supressed (S) due to a privacy clause.

Table 4.3: Exogeneity Check: without students born in May or June

		Main Model	w/o May & June
Age m	*Coefficients*	-0.479***	-1.235***
	Standard errors	0.004	0.006
Age m2	*Coefficients*	0.672***	1.175***
	Standard errors	0.003	0.004
Female	*Coefficients*	0.001	0.001
	Standard errors	0.001	0.001
Ethnicity: Maori	*Coefficients*	-0.003***	-0.004***
	Standard errors	0.001	0.001
Ethnicity: Asian	*Coefficients*	0.000	-0.001
	Standard errors	0.001	0.001
Ethnicity: Australian	*Coefficients*	-0.002	-0.004
	Standard errors	0.003	0.004
Ethnicity: European	*Coefficients*	0.003	0.002
	Standard errors	0.001	0.001
Ethnicity: Pacific People	*Coefficients*	0.001	0.001
	Standard errors	0.001	0.001
Number of observations		411,765	344,214

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.4: Effect of maximum schooling: without students born in May or June

		Main Model	w/o May & June
NCEA1	*Coefficients*	0.021***	0.012**
	Standard errors	(0.003)	(0.003)
	Percentage change	2.4%	1.3%
NCEA2	*Coefficients*	0.037***	0.028***
	Standard errors	(0.004)	(0.004)
	Percentage change	4.4%	3.3%
NCEA3	*Coefficients*	0.039***	0.034***
	Standard errors	(0.004)	(0.005)
	Percentage change	6.2%	5.4%
UE	*Coefficients*	0.021***	0.022***
	Standard errors	(0.004)	(0.004)
	Percentage change	5.0%	5.2%
Number of observations	NCEA	331,320	276,933
	UE	411,765	344,214

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.5: Exogeneity Check: different date of birth assumptions

		Main Model	1st of each month	15th of each month	Last of each month
Age m	*Coefficients*	-0.479***	-0.988***	-0.535***	-0.049***
	Standard errors	0.004	0.002	0.003	0.004
Age m2	*Coefficients*	0.672***	1.010***	0.710***	0.382***
	Standard errors	0.003	0.002	0.002	0.003
Female	*Coefficients*	0.001	0.000	0.001*	0.002**
	Standard errors	0.001	0.001	0.001	0.001
Ethnicity: Maori	*Coefficients*	-0.003***	-0.002**	-0.004***	-0.003***
	Standard errors	0.001	0.001	0.001	0.001
Ethnicity: Asian	*Coefficients*	0.000	0.000	0.000	0.001
	Standard errors	0.001	0.001	0.001	0.001
Ethnicity: Australian	*Coefficients*	-0.002	-0.004	-0.006*	-0.002
	Standard errors	0.003	0.003	0.003	0.003
Ethnicity: European	*Coefficients*	0.003	0.002	0.003*	0.004**
	Standard errors	0.001	0.001	0.001	0.001
Ethnicity: Pacific People	*Coefficients*	0.001	0.000	0.001	0.003*
	Standard errors	0.001	0.001	0.001	0.001
Number of observations		411,765	411,765	411,765	411,765

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.6: Effect of maximum schooling: different date of birth assumptions

		Main Model	**1st of each month**	**15th of each month**	**Last of each month**
NCEA1	*Coefficients*	0.021***	0.014***	0.023***	0.029***
	Standard errors	(0.003)	(0.003)	(0.003)	(0.003)
	Percentage change	2.4%	1.6%	2.6%	3.3%
NCEA2	*Coefficients*	0.037***	0.027***	0.039***	0.044***
	Standard errors	(0.004)	(0.004)	(0.004)	(0.004)
	Percentage change	4.4%	3.2%	4.6%	5.2%
NCEA3	*Coefficients*	0.039***	0.025***	0.040***	0.045***
	Standard errors	(0.004)	(0.005)	(0.004)	(0.005)
	Percentage change	6.2%	4.0%	6.4%	7.2%
UE	*Coefficients*	0.021***	0.008	0.022***	0.029***
	Standard errors	(0.004)	(0.005)	(0.004)	(0.004)
	Percentage change	5.0%	1.9%	5.2%	6.8%
Number of observations	NCEA	331,320	331,320	331,320	331,320
	UE	411,765	411,765	411,765	411,765

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.7: Effect of maximum schooling: marginal effects of ordered logit and ordered probit

		Ordered Logit	Ordered Probit
NCEA0	*dy/dx*	-0.017***	-0.018***
	Standard errors	0.002	0.002
NCEA1	*dy/dx*	-0.007***	-0.007***
	Standard errors	0.001	0.001
NCEA2	*dy/dx*	-0.023***	-0.017***
	Standard errors	0.005	0.004
NCEA3	*dy/dx*	0.047***	0.042***
	Standard errors	0.007	0.006
Number of observations		331,320	331,320

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively. The marginal effects are evaluated at mean values.

Table 4.8: Population for placebo group 1: migrants and refugees

Placebo Group 1				
	Domestic	**IFP**	**Exchange**	**Total**
Migrant	32,169	210	63	**32,379**
Refugee	2,655	S	S	**2,655**
NZ Born	20,613	S	S	**20,613**
Missing	391,290	13,959	510	**405,249**
Total	**446,727**	**14,169**	**573**	**460,896**

Note: Grey highlighted is the population for the main analysis and blue highlighted is the population for the placebo 1 (migrants and refugee students) analysis.

Table 4.9: Population for placebo group 2: international fee paying students

Placebo Group 2				
	Domestic	**IFP**	**Exchange**	**Total**
Migrant	32,169	210	63	**32,379**
Refugee	2,655	S	S	**2,655**
NZ Born	20,613	S	S	**20,613**
Missing	391,290	13,959	510	**405,249**
Total	**446,727**	**14,169**	**573**	**460,896**

Note: Grey highlighted is the population for the main analysis and blue highlighted is the population for the placebo 2 (IFP students) analysis.

Table 4.10: Ethnic composition of population in placebo tests

ETHNICITY	MAIN		PLACEBO GROUP 1 (Migrants and Refugees)		PLACEBO GROUP 2 (International Fee Paying)	
	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE
New Zealand European	248,670	60.39%	726	2.07%	-	-
British and Irish	9,321	2.26%	18	0.05%	33	0.23%
Dutch	489	0.12%	168	0.48%	18	0.13%
Greek	135	0.03%	9	0.03%	-	-
Polish	57	0.01%	36	0.10%	-	-
South Slav	93	0.02%	108	0.31%	-	-
Italian	93	0.02%	24	0.07%	96	0.68%
German	348	0.08%	147	0.42%	486	3.43%
Australian	2,511	0.61%	18	0.05%	-	-
Other European	6,897	1.67%	1,284	3.67%	360	2.54%
Māori	80,163	19.47%	63	0.18%	-	-
Samoan	14,568	3.54%	4,485	12.81%	60	0.42%
Cook Islands Māori	4,854	1.18%	438	1.25%	-	-
Tongan	6,522	1.58%	1,716	4.90%	114	0.81%
Niuean	1,674	0.41%	66	0.19%	-	-
Tokelauan	606	0.15%	141	0.40%	-	-
Fijian	1,608	0.39%	1,074	3.07%	234	1.65%
Other Pacific Peoples	1,188	0.29%	603	1.72%	114	0.81%
Filipino	2,502	0.61%	2,457	7.02%	18	0.13%
Cambodian	435	0.11%	369	1.05%	45	0.32%
Vietnamese	327	0.08%	351	1.00%	549	3.88%
Other Southeast Asian	1,005	0.24%	906	2.59%	513	3.62%
Chinese	7,521	1.83%	5,070	14.48%	6,675	47.14%
Indian	8,160	1.98%	5,244	14.98%	114	0.81%
Sri Lankan	618	0.15%	438	1.25%	6	0.04%
Japanese	450	0.11%	222	0.63%	1,098	7.75%
Korean	1,770	0.43%	2,898	8.28%	2,193	15.49%
Other Asian	1,812	0.44%	1,566	4.47%	741	5.23%
Middle Eastern	813	0.20%	1,623	4.64%	252	1.78%
Latin American	267	0.06%	426	1.22%	348	2.46%
African	3,330	0.81%	1,485	4.24%	15	0.11%
Other Ethnicity	1,890	0.46%	672	1.92%	51	0.36%
Do not Know	264	0.06%	36	0.10%	6	0.04%
Not Stated	804	0.20%	123	0.35%	21	0.15%
Total	**411,765**	**100.00%**	**35,010**	**100.00%**	**14,160**	**100.00%**

Note: All figures have been randomly rounded to base 3 (RR3) – the number is randomly rounded to either the nearest base above or below the number – following the Stats NZ privacy requirement.

Table 4.11: School decile composition of population in placebo tests

SCHOOL DECILE	MAIN		PLACEBO GROUP 1 (Migrants and Refugees)		PLACEBO GROUP 2 (International Fee Paying)	
	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE
1	21,534	5.23%	4,338	12.39%	102	0.72%
2	21,996	5.34%	2,673	7.63%	117	0.83%
3	32,139	7.81%	4,065	11.61%	666	4.70%
4	39,372	9.56%	3,852	11.00%	672	4.75%
5	32,046	7.78%	1,449	4.14%	402	2.84%
6	58,278	14.15%	3,450	9.85%	1,176	8.31%
7	47,859	11.62%	3,564	10.18%	4,998	35.30%
8	45,036	10.94%	2,706	7.73%	1,467	10.36%
9	49,899	12.12%	5,538	15.82%	2,514	17.75%
10	41,637	10.11%	2,727	7.79%	1,773	12.52%
Missing	21,969	5.34%	648	1.85%	273	1.93%
Total	**411,765**	**100.00%**	**35,010**	**100.00%**	**14,160**	**100.00%**

Note: All figures have been randomly rounded to base 3 (RR3) – the number is randomly rounded to either the nearest base above or below the number – following the Stats NZ privacy requirement.

Table 4.12: NCEA achievement in placebo populations

NCEA	MAIN		PLACEBO GROUP 1 (Migrants and Refugees)		PLACEBO GROUP 2 (International Fee Paying)	
	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE
0	36,483	8.86%	2,613	7.46%	4,845	34.22%
1	16,545	4.02%	642	1.83%	171	1.21%
2	69,927	16.98%	4,443	12.69%	1,047	7.39%
3	208,362	50.60%	19,530	55.78%	5,937	41.93%
Missing	80,448	19.54%	7,782	22.23%	2,160	15.25%
Total	**411,765**	**100.00%**	**35,010**	**100.00%**	**14,160**	**100.00%**

Note: All figures have been randomly rounded to base 3 (RR3) – the number is randomly rounded to either the nearest base above or below the number – following the Stats NZ privacy requirement.

Table 4.13: UE achievement in placebo populations

UE	MAIN		PLACEBO GROUP 1 (Migrants and Refugees)		PLACEBO GROUP 2 (International Fee Paying)	
	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE	FREQ.	PERCENTAGE
0	236,916	57.54%	18,894	53.97%	8,724	61.61%
1	174,849	42.46%	16,116	46.03%	5,436	38.39%
Total	**411,765**	**100.00%**	**35,010**	**100.00%**	**14,160**	**100.00%**

Note: All figures have been randomly rounded to base 3 (RR3) – the number is randomly rounded to either the nearest base above or below the number – following the Stats NZ privacy requirement.

Table 4.14: Exogeneity Check: placebo tests

		Main Model	Placebo Group 1 (Migrants and Refugees)	Placebo Group 2 (International Fee Paying)
Age m	*Coefficients*	-0.481***	-0.451***	-0.421***
	Standard errors	0.004	0.012	0.028
Age m2	*Coefficients*	0.673***	0.652***	0.633***
	Standard errors	0.002	0.008	0.019
Female	*Coefficients*	0.001	-0.002	0.001
	Standard errors	0.001	0.002	0.003
Ethnicity: Maori	*Coefficients*	-0.003***	0.007	-
	Standard errors	0.001	0.020	-
Ethnicity: Asian	*Coefficients*	0.000	-0.001	0.040
	Standard errors	0.001	0.007	0.042
Ethnicity: Australian	*Coefficients*	-0.003	0.018	-
	Standard errors	0.003	0.041	-
Ethnicity: European	*Coefficients*	0.003	-0.002	0.052
	Standard errors	0.001	0.008	0.042
Ethnicity: Pacific People	*Coefficients*	0.001	0.003	0.056
	Standard errors	0.001	0.007	0.041
Number of observations		391,368	35,010	14,157

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.15: Effect of maximum schooling: placebo tests

		Main Model	Placebo Group 1 (Migrants and Refugees)	Placebo Group 2 (International Fee Paying)
NCEA1	*Coefficients*	0.021***	0.002	0.023
	Standard errors	(0.003)	(0.010)	(0.016)
	Percentage change	2.4%	0.0%	3.9%
NCEA2	*Coefficients*	0.037***	0.000	0.025
	Standard errors	(0.004)	(0.010)	(0.017)
	Percentage change	4.4%	0.0%	4.3%
NCEA3	*Coefficients*	0.039***	-0.001	0.043*
	Standard errors	(0.004)	(0.015)	(0.021)
	Percentage change	6.2%	-0.1%	8.7%
UE	*Coefficients*	0.021***	-0.002	0.053**
	Standard errors	(0.004)	(0.015)	(0.020)
	Percentage change	5.0%	-0.4%	13.8%
Number of observations	NCEA	331,320	27,225	11,997
	UE	411,765	35,010	14,157

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.16: Exogeneity Check: by gender

		Main Model	Female	Male
Age m	*Coefficients*	-0.479***	-0.474***	-0.484***
	Standard errors	0.004	0.005	0.005
Age m2	*Coefficients*	0.672***	0.668***	0.675***
	Standard errors	0.003	0.003	0.004
Female	*Coefficients*	0.001	0.000	0.000
	Standard errors	0.001	(omitted)	(omitted)
Ethnicity: Maori	*Coefficients*	-0.003***	-0.004***	-0.003**
	Standard errors	0.001	0.001	0.001
Ethnicity: Asian	*Coefficients*	0.000	-0.001	0.001
	Standard errors	0.001	0.002	0.001
Ethnicity: Australian	*Coefficients*	-0.002	0.001*	-0.006
	Standard errors	0.003	0.005	0.005
Ethnicity: European	*Coefficients*	0.003	0.005*	0.002
	Standard errors	0.001	0.002	0.002
Ethnicity: Pacific People	*Coefficients*	0.001	0.001	0.002
	Standard errors	0.001	0.002	0.002
Number of observations		411,765	201,522	210,246

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.17: Effect of maximum schooling: by gender

		Main Model	Female	Male
NCEA1	*Coefficients*	0.021***	0.018***	0.025***
	Standard errors	(0.003)	(0.004)	(0.005)
	Percentage change	2.4%	2.0%	2.9%
NCEA2	*Coefficients*	0.037***	0.027***	0.047***
	Standard errors	(0.004)	(0.006)	(0.006)
	Percentage change	4.4%	3.1%	5.8%
NCEA3	*Coefficients*	0.039***	0.037***	0.041***
	Standard errors	(0.004)	(0.006)	(0.007)
	Percentage change	6.2%	5.3%	7.4%
UE	*Coefficients*	0.021***	0.020**	0.022***
	Standard errors	(0.004)	(0.006)	(0.006)
	Percentage change	5.0%	4.0%	6.2%
Number of observations	NCEA	331,320	167,862	163,458
	UE	411,765	201,522	210,246

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.18: Exogeneity Check: by ethnicity

		Main Model	NZ European	Māori	Asian
Age m	*Coefficients*	-0.479***	-0.482***	-0.482***	-0.506***
	Standard errors	0.004	0.005	0.008	0.013
Age m2	*Coefficients*	0.672***	0.674***	0.674***	0.690***
	Standard errors	0.003	0.003	0.006	0.008
Female	*Coefficients*	0.001	0.001	0.001	-0.001
	Standard errors	0.001	0.001	0.001	0.002
Number of observations		411,765	248,667	80,163	24,606

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.19: Effect of maximum schooling: by ethnicity

		Main Model	NZ European	Māori	Asian
NCEA1	*Coefficients*	0.021***	0.015***	0.056***	-0.002
	Standard errors	(0.003)	(0.003)	(0.010)	(0.006)
	Percentage change	2.4%	1.6%	7.4%	-0.2%
NCEA2	*Coefficients*	0.037***	0.037***	0.060***	0.007
	Standard errors	(0.004)	(0.005)	(0.011)	(0.007)
	Percentage change	4.4%	4.2%	8.9%	0.7%
NCEA3	*Coefficients*	0.039***	0.056***	0.027**	0.010
	Standard errors	(0.004)	(0.006)	(0.010)	(0.013)
	Percentage change	6.2%	8.5%	6.7%	1.1%
UE	*Coefficients*	0.021***	0.033***	0.022**	0.013
	Standard errors	(0.004)	(0.006)	(0.008)	(0.015)
	Percentage change	5.0%	6.9%	11.9%	1.7%
Number of observations	NCEA	331,320	206,310	58,737	21,867
	UE	411,765	248,667	80,163	24,606

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.20: Exogeneity Check: by school decile group

		Main Model	Decile group 1 (School deciles 1–4)	Decile group 2 (School deciles 5–7)	Decile group 3 (School deciles 8–10)
Age m	*Coefficients*	-0.479***	-0.471***	-0.485***	-0.487***
	Standard errors	0.004	0.007	0.006	0.006
Age m^2	*Coefficients*	0.672***	0.666***	0.676***	0.677***
	Standard errors	0.003	0.005	0.004	0.004
Female	*Coefficients*	0.001	0.001	0.001	0.001
	Standard errors	0.001	0.001	0.001	0.001
Ethnicity: Maori	*Coefficients*	-0.003***	-0.002*	-0.005***	-0.003
	Standard errors	0.001	0.001	0.001	0.002
Ethnicity: Asian	*Coefficients*	0.000	-0.001	0.002	-0.002
	Standard errors	0.001	0.002	0.002	0.001
Ethnicity: Australian	*Coefficients*	-0.002	0.005	-0.011	0.001
	Standard errors	0.003	0.008	0.006	0.005
Ethnicity: European	*Coefficients*	0.003	-0.001	0.006*	0.003
	Standard errors	0.001	0.005	0.002	0.002
Ethnicity: Pacific People	*Coefficients*	0.001	0.000	0.004	0.002
	Standard errors	0.001	0.002	0.003	0.003
Number of observations		411,765	115,038	138,180	136,575

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

Table 4.21: Effect of maximum schooling: by school decile group

		Main Model	Decile group 1 (School deciles 1–4)	Decile group 2 (School deciles 5–7)	Decile group 3 (School deciles 8–10)
NCEA1	*Coefficients*	0.021***	0.028**	0.025***	0.011**
	Standard errors	(0.003)	(0.008)	(0.005)	(0.003)
	Percentage change	2.4%	3.4%	2.8%	1.1%
NCEA2	*Coefficients*	0.037***	0.034***	0.046***	0.026***
	Standard errors	(0.004)	(0.008)	(0.007)	(0.005)
	Percentage change	4.4%	4.5%	5.4%	2.8%
NCEA3	*Coefficients*	0.039***	0.016	0.059***	0.036***
	Standard errors	(0.004)	(0.008)	(0.009)	(0.007)
	Percentage change	6.2%	3.2%	9.8%	4.5%
UE	*Coefficients*	0.021***	0.011	0.039***	0.014
	Standard errors	(0.004)	(0.007)	(0.008)	(0.008)
	Percentage change	5.0%	4.3%	9.9%	2.2%
Number of observations	NCEA	331,320	85,314	110,070	118,959
	UE	411,765	115,038	138,180	136,575

Note: All regressions also include thirteen year of birth dummies and their interactions with eighteen region dummies; and school fixed effects. Standard errors are robust and clustered by school. *, **, and *** indicate statistical significance at 95%, 99%, and 99.9% confidence levels respectively.

CHAPTER 5. CONCLUSION

In this thesis, I undertake two different studies on the returns to early schooling: 1) an analysis of short-term educational outcomes via a replication of a Dutch paper by Leuven et al (2010) and 2) an extensive original study examining the effects of early schooling on long-term achievement using confidentialised micro-data from New Zealand. The unique structure of schooling in the Netherlands and New Zealand allows me to investigate the causal effects of schooling on later educational outcomes under the exogeneity of all other factors.

My replication of Leuven et al. (2010), in Chapter 2, examines the effects of early formal education starting on a child's 4th birthday only about two years later, around the age of six. Both the original study and the replication find positive effects and the replication in general endorses the conclusions of Leuven et al. (2010), but with some noteworthy differences. For disadvantaged children, the replication finds beneficial effects of early schooling that are slightly smaller than those in the original study. On the other hand, for non-disadvantaged children, the replication finds strong, positive and large effects, contrary to Leuven et al's conclusions.

Chapter 3 uses data from New Zealand to study longer-term consequences. In particular, it focuses on effects of early schooling after 10-13 years, around the age of 15-18. A qualitative comparison with Chapter 2 is made possible by the fact that New Zealand and the Netherlands have a similar primary school entrance policy, where children typically start school right on their 5th and 4th birthday, respectively. Both chapters find large beneficial effects of early schooling and Chapter 3 further suggests that, if anything, such effects become larger over time as students move to more advanced learning.

Chapter 4 consists of a series of robustness, falsification and homogeneity tests, which serve to bolster the credibility of the large effects observed in Chapter 3. The various robustness checks, such as removing the most 'noisy' observations from the sample and employing alternative estimators, suggest that my results are robust. Similarly, my falsification checks tentatively indicate that the effects in my main analysis are indeed driven by experiences in early schooling. Finally, Chapter 4 also suggests that the observed beneficial effects are the strongest among male, Māori, and decile 5-7 children.

Taken together, the analyses presented in this thesis indicates that 1) early schooling experiences are indeed an important part of cognitive development and have long-term effects on academic achievement and 2) the timing of entry into early formal education can be an important decision for parents and schools – where ‘holding back’ a child may do more for his/her academic future than speedy transition into higher grades.

REFERENCES

Ali, A. I., & Menclova, A. (2018). *Returns to Early Formal Education: A Replication of Leuven et al. (2010).* Working Papers in Economics 18/19, University of Canterbury, Department of Economics and Finance.

Anae, M., Anderson, H., Benseman, J., & Coxon, E. (2002). *Pacific peoples tertiary education: Issues of participation.* Ministry of Education. Wellington: Auckland Uniservices Limited.

Angrist, J. D., & Krueger, A. B. (1992, June). The Effect of Age at School Entry on Educational Attainment: An Application ofInstrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association, 87*(418), 328-336.

Barnsley, R. H., & Thompson, A. H. (1988). Birthdate and success in minor hockey: The key to the NHL. *Canadian Journal of Behavioural Science, 20*(2), 167-176.

Barnsley, R. H., Thompson, A. H., & Barnsley, E. P. (1985). Hockey success and birthdate: The relative age effect. *Journal of the Canadian Association of Health, Physical Education and Recreation, 51*.

Boucher, J. L., & Mutimer, B. T. (1994). The relative age phenomenon in sport: A replication and extension with ice-hockey players. *Research Quarterly for Exercise and Sport, 65*(4), 377-381.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development, 60*, 1239-149.

Caudill, S. B. (1988). Practitioners corner: An advantage of the linear probability model over probit or logit. *Oxford Bulletin of Economics and Statistics, 50*(4), 425-527.

Clark, J., Kim, B., Poulton, R., & Milne, B. (2006). The role of low expectations in health and education investment and hazardous consumption. *Canadian Journal of Economics/Revue canadienne d'économique, 39*(4), 1151-1172.

Cliffordson, C., & Gustafsson, J.-E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence, 36*(2), 143-152.

Cobley, S., Baker, J., Wattie, N., & McKenna, J. (2009). Annual Age-Grouping and Athlete Development: A Meta-Analytical Review of Relative Age Effects in Sport. *Sports Medicine, 39*(3), 235-256.

Crawford, C., Dearden, L., & Greaves, E. (2013). *When You Are Born Matters: The Impact of Date of Birth on Educational Outcomes in England.* London: Institute for Fiscal Studies (IFS).

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the Evidence on Life Cycle Skill. *Handbook of the Economics of Education, 1*, 697-812.

Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review, 85, No. 3,* 341-364.

Daniel, T. E., & Janssen, C. T. (1987). More on the relative age effect. *Canadian Association for Health, Physical Education, and Recreation, 53*, 21-24.

Datar, A. (2004). Does delaying kindergarten entrance give children a head start? *Economics of education review, 25*, 43-62.

DiPasquale, G. W., Moule,, A., & Flewelling, R. (1980). The Birthdate Effect. *Journal of Learning Disabilities, 13*, 234-248.

Dubin, J. A., & Rivers, D. (1989). Selection bias in linear regression, logit and probit models. *Sociological Methods & Research, 18*(2-3), 360-390.

Earle, D. (2010). How can tertiary education deliver better value to the economy? *Tertiary Sector Performance Analysis and Reporting, Strategy and System Performance.* Ministry of Education.

Edwards, J. R. (1974). Characteristics of disadvantaged children. *The Irish journal of education, 1*, 49-61.

Friedman, J. (2012, July). *Whether to probit or to probe it: in defense of the Linear Probability Model.* Retrieved from World Bank Blogs: http://blogs.worldbank.org/impactevaluations/whether-to-probit-or-to-probe-it-in-defense-of-the-linear-probability-model

Friedman, J., & Schady, N. (2013). How many infants likely died in Africa as a result of the 2008–2009 global financial crisis? *Health Economics, 22(5),* 611-622.

Fumarco, L., Gibbs, B. G., Jarvis, J. A., & Rossi, G. (2017). The relative age effect reversal among the National Hockey League elite. *Plos One, 12*(8).

Grondin, S., Deshaies, P., & Nault, L. (1984). Trimestres de naissance et participation au hockey et au volleyball. *La Revue Quebecoise de l'Activite Physique, 2*, 97-103.

Hanushek, E. A., & Wößmann, L. (2007). *Education Quality and Economic Growth.* The World Bank.

Heckman, J. J. (2000, March). Policies to Foster Human Capital. *Research in Economics, 5*(1), 3-56.

Hippel, P. V. (2015, July). *Linear vs. Logistic Probability Models: Which is Better, and When?* Retrieved from Statistical Horizons: https://statisticalhorizons.com/linear-vs-logistic

Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal, 28*(4), 331-343.

Kautz, T., Heckman, J. J., & Diris, R. (2014). *Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success.* The National Bureau of Economic Research.

Kawaguchi, D. (2011). Actual age at school entry, educational outcomes, and earnings. *Journal of the Japanese and International Economies, 25*(2), 64-80.

Leeuwen, B. v., Thijs, A., & Zandbergen, M. (2009). *Inclusive education in the Netherlands.* SLO - National institute for curriculum development.

Leuven, E., Lindahl, M., Oosterbeek, H., & Webbink, D. (2010, June). Expanding schooling opportunities for 4-year-olds. *Economics of Education Review, 29*(3), 319-328.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *Review of Economic Studies, 72*, 189-221.

Lubotsky, D., & Kaestner, R. (2016). Do 'Skills Beget Skills'? Evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood. *Economics of Education Review, 53*, 194-206.

Ministry of Education. (2015). New Zealand Government.

Musch, J., & Grondin, S. (2001). Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sport. *Developmental Review, 21*, 147–167.

Nakhid, C. (2003). " Intercultural" Perceptions, Academic Achievement, and the Identifying Process of Pacific Islands Students in New Zealand Schools. *Journal of Negro Education*, 297-317.

New Zealand Qualifications Authority. (2013/14). New Zealand Government.

Programme for International Student Assessment. (2016).

Shui, M. (2017). *What grades do I need to get into a NZ university?* MyTuition.

Strøm, B. (2004). Student achievement and birthday effects. *Schooling and Human Capital in the Global Economy: Revisiting the Equity-Efficiency Quandary.* Munich: CESifo-Harvard University/PEPG Conference.

The Treasury - Kaitohutohu Kaupapa Rawa. (2015). New Zealand Government.

Tofi, T., Flett, R., & Timutimu-Thorpe, H. (1996). Problems faced by pacific island students at university in New Zealand: some effects on academic performance and psychological wellbeing. *New Zealand Journal of Educational Studies, 31*(1), 51-59.

UNESCO Institute for Statistics. (2012). The World Bank.

UNESCO Institute for Statistics. (2016). The World Bank.

Warren, C. W., Levin, M., & Tyler, C. (1986). Season of Birth and Academic Achievement. *Educational and Psychological Research, 6*, 111-123.

Zhang, S., Zhong, R., & Zhang, J. (2017). School starting age and academic achievement: Evidence from China's junior high schools. *China Economic Review, 44*, 343-354.